Linear Optimal Control, Filtering, and Rational Expectations

by

Thomas J. Sargent

Hoover Institution
Stanford University

Revised
November 1988

Table of Contents

Chapter 1

Classical Control and Prediction Theory

1. Introduction

A theme that will recur throughout these pages is that there is an intimate connection between two superficially different classes of problems: the class of linear-quadratic optimal control problems, and the class of linear least squares predicition and filtering problems. The classes of problems are connected in the sense that to solve each, essentially the same mathematics is used. This connection, which is often termed "duality," is present whether one uses "classical" or "recursive" solution procedures.[1] It is worthwhile to exhibit this interconnection early on. We do this here in the simple context of a pair of univariate examples.

2. An Infinite Horizon Control Problem

Consider the discrete time control problem, to maximize

$$\lim_{N\to\infty} \sum_{t=0}^{N} \beta^t \left\{a_t y_t - \frac{1}{2} h y_t^2 - \frac{1}{2}\left[d(L) y_t\right]^2\right\},\ h > 0,\ 0 < \beta < 1 \tag{1.1}$$

where $d(L) = d_0 + d_1 L + \ldots + d_m L^m$, L is the lag operator, $\{a_t,\ t = 0,\ 1, \ldots\}$ is a sequence of exponential order less than $\beta^{-1/2}$, and β is the discount factor. The maximization in (1.1) is subject to the initial conditions for $y_{-1}, y_{-2} \ldots, y_{-m}$. The maximization is over infinite sequences for $y_t, t = 0, 1, \ldots$. Simple examples of this problem for factor demand, economic growth, and government policy problems are given in Sargent [ch. 9].

We first study a finite N version of the problem. Our approach will be to study the limit of the solution of the finite N problem. This will require being careful, as indicated below, because the limits as N approaches infinity of the necessary and sufficient conditions for maximizing finite N versions of (1.1) are not sufficient for maximizing (1.1).

We begin by fixing $N > m$, differentiating the finite version of (1.1) with respect to $y_0, y_1, \ldots, y_N$, and then setting these derivatives to zero. For $t = 0, \ldots, N - m$ these first

[1] By "classical" procedures, we mean solution of the control problem via discrete time variational methods, and solution of the prediction problem by the Wiener-Kolmogorov method. By "recursive" procedures, we mean solution of the control problem by iterating on the matrix Riccati difference equation, and solution of the prediction problem via the Kalman filter.

order necessary conditions are the Euler equations. For $t = N-m+1, \ldots, N$, the first order conditions are a set of terminal conditions.

In carrying out this differentiation, the only problematic term is

$$\sum_{t=0}^{\infty} \beta^t \left[d(L)y_t\right]^2.$$

Consider the term

$$\begin{aligned}
L &= \sum_{t-0}^{N} \beta^t \left[d(L)y_t\right]\left[d(L)y_t\right] \\
&= \sum_{t=0}^{N} \beta^t \left(d_0\, y_t + d_1\, y_{t-1} + \cdots + d_m\, y_{t-m}\right)\left(d_0\, y_t + d_1\, y_{t-1} + \cdots + d_m\, y_{t-m}\right).
\end{aligned}$$

Differentiating L with respect to y_t for $t = 0,\ 1,\ \ldots,\ N-m$ gives

$$\begin{aligned}
\frac{\partial L}{\partial y_t} &= \beta^t\, d_0\, d(L)y_t + \beta^{t+1}\, d_1\, d(L)y_{t+1} + \cdots + \beta^{t+m}\, d_m\, d(L)y_{t+m} \\
&\quad + \beta^t\, d_0\, d(L)y_t + \beta^{t+1}\, d_1\, d(L)y_{t+1} + \cdots + \beta^{t+m}\, d_m\, d(L)y_{t+m} \\
&= 2\beta^t \left(d_0 + d_1\, \beta L^{-1} + d_2\, \beta^2\, L^{-2} + \cdots + d_m\, \beta^m\, L^{-m}\right) d(L)y_t.
\end{aligned}$$

So we have

$$\frac{\partial L}{\partial y_t} = 2\beta^t\, d\left(\beta L^{-1}\right) d(L)y_t. \tag{1.2}$$

Differentiating L with respect to y_t for $t = N-m+1, \ldots, N$ gives

$$\begin{aligned}
\frac{\partial L}{\partial y_N} &= 2\beta^N\, d_0\, d(L)y_N \\
\frac{\partial L}{\partial y_{N-1}} &= 2\beta^{N-1}\left[d_0 + \beta\, d_1\, L^{-1}\right] d(L)y_{N-1} \\
&\ \ \vdots \\
\frac{\partial L}{\partial y_{N-m+1}} &= 2\beta^{N-m+1}\left[d_0 + \beta L^{-1}\, d_1 + \cdots + \beta^{m-1}\, L^{-m+1}\, d_{m-1}\right] d(L)y_{N-m+1}.
\end{aligned} \tag{1.3}$$

The derivatives (1.2) and (1.3) are the keys to obtaining the Euler equations and the transversality conditions, respectively.

Differentiating (1.1) with respect to y_t for $t = 0, \ldots, N-m$ gives the Euler equations

$$\left[h + d\left(\beta L^{-1}\right) d(L)\right] y_t = a_t,\ \ t = 0,\ 1,\ \ldots, N-m. \tag{1.4}$$

Differentiating (1.1) with respect to y_t for $t = N - m + 1, \ldots, N$ gives the terminal conditions

(1.5)

$$\beta^N (a_N - hy_N - d_0\, d(L) y_N) = 0$$

$$\beta^{N-1}\left(a_{N-1} - hy_{N-1} - \left(d_0 + \beta\, d_1\, L^{-1}\, d(L)\right) y_{N-1}\right) = 0$$

$$\vdots$$

$$\beta^{N-m+1}\left(a_{N-m+1} - hy_{N-m+1} - (d_0 + \beta L^{-1} d_1 + \cdots + \beta^{m-1} L^{-m+1} d_{m-1}) d(L) y_{N-m+1}\right) = 0.$$

In the finite N problem, we have to solve the Euler equation (1.4), which is a $2m^{\text{th}}$ order linear difference equation, subject to the m initial conditions $y_1, \ldots, y_m$ and the m terminal conditions (1.5). These conditions uniquely determine the correct solution in the finite N problem. That is, for the finite N problem, conditions (1.4) and (1.5) are necessary and sufficient for a maximum. In Section 6 below, we shall briefly describe representations of the solution using matrix methods.

For the infinite horizon problem, we propose to discover first-order necessary conditions by taking the limits of (1.4) and (1.5) as N goes to infinity. This approach is valid, and the limits of (1.4) and (1.5) as N approaches infinity are first-order necessary conditions for a maximum. However, for the infinite horizon problem with $\beta < 1$, the limits of (1.4) and (1.5) are, in general, not sufficient conditions for a maximum. That is, the limits of (1.5) do not provide enough information uniquely to determine the solution of the Euler equation (1.4) that maximizes (1.1). As it turns out, and as we shall see below, a side condition on the path of y_t that together with (1.4) is sufficient for an optimum is

$$\sum_{t=0}^{\infty} \beta^t h y_t^2 < +\infty. \tag{1.6}$$

All paths that satisfy the Euler equations, except the one that we shall select below, violate this condition and, therefore, evidently lead to (much) lower values of the criterion function (1.1) than does the optimal path selected by the solution procedure below.

Consider the characteristic equation for the Euler equation

$$[h + d(\beta z^{-1})\, d(z)] = 0. \tag{1.7}$$

Notice that if $\tilde{z}$ is a root of equation (1.7), then so is $\beta \tilde{z}^{-1}$. Thus, the roots of (1.7) come in "β-reciprocal" pairs. If $\beta = 1$, the roots come in reciprocal pairs. Assume that the roots of (7) are distinct.[2] Let the roots be, in descending order according to their moduli, $z_1, z_2, \cdots, z_m, z_{m+1}, \cdots, z_{2m}$, so that $|z_1| > |z_2| > \ldots |z_m| > |z_{m+1}| > \ldots > |z_{2m}|$. From the pairs property and the assumption of distinct roots, it follows that $|z_j| > \sqrt{\beta}$ for $j < m$ and $|z_j| < \sqrt{\beta}$ for $j > m$. It also follows that $z_{2m-j} = \beta z_{j+1}^{-1}, j = 0, 1, \ldots, m-1$. Therefore, the characteristic polynomial on the left side of (1.7) can be expressed as[3]

(1.8)

$$\begin{aligned}\left[h + d(\beta z^{-1})d(z)\right] &= z^{-m} z_0 (z - z_1) \cdots (z - z_m)(z - z_{m+1}) \cdots (z - z_{2m}) \\ &= z^{-m} z_0 (z - z_1)(z - z_2) \cdots (z - z_m)(z - \beta z_m^{-1}) \cdots (z - \beta z_2^{-1})(z - \beta z_1^{-1}),\end{aligned}$$

where z_0 is a constant. In (1.8), we substitute $(z - z_j) = -z_j(1 - \frac{1}{z_j} z)$ and $(z - \beta z_j^{-1}) = z(1 - \frac{\beta}{z_j} z^{-1})$ for$j = 1, \ldots, m$ to get

$$\left[h + d(\beta z^{-1})d(z)\right] = (-1)^m (z_0 z_1 \cdots z_m)(1 - \frac{1}{z_1} z) \cdots (1 - \frac{1}{z_m} z)(1 - \frac{1}{z_1} \beta z^{-1}) \cdots (1 - \frac{1}{z_m} \beta z^{-1}).$$

Now define $c(z) = \sum_{j=0}^{m} c_j z^j$ as

$$c(z) = \left[(-1)^m z_0 z_1 \cdots z_m\right]^{1/2} (1 - \frac{z}{z_1})(1 - \frac{z}{z_2}) \cdots (1 - \frac{z}{z_m}). \tag{1.9}$$

Then notice that (1.8) can be written

$$h + d(\beta z^{-1})\, d(z) = c(\beta z^{-1})\, c(z). \tag{1.10}$$

It is useful to write (1.9) as

$$c(z) = c_0 (1 - \lambda_1 z) \ldots (1 - \lambda_m z) \tag{1.11}$$

where

$$c_0 = \left[(-1)^m z_0 z_1 \cdots z_m\right]^{1/2}; \ \lambda_j = \frac{1}{z_j}, \ j = 1, \ldots, m.$$

---

[2] We make this assumption mainly for convenience. The development below can readily be modified to accommodate repeated roots of (1.7), using Gabel and Roberts [ ], Churchill [ ], or Sargent [, ch. 9]. From a practical point of view, the assumption of distinct roots is not very restrictive since systems with repeated roots can be approximated arbitrarily well by systems with distinct roots.

[3] These expressions are correct even if there is a repeated root of (1.7) at zero, that is, even if $z_m = z_{m+1} = \sqrt{\beta}$. Most, but not all, of the subsequent results on prediction and control go through if $z_m = \sqrt{\beta}$. The optimal feedback law (1.14) or (1.15) holds with $z_m^{-1} = \lambda_m = \beta^{-1/2}$ if we restrict the a_t sequence to be of exponential order less than $\frac{1}{\sqrt{\beta}}$. The Weiner-Kolmogorov formula (1.27) is not appropriate if $z_m = 1$, because $c(L)^{-1}$ does not exist. However, a modified version of the Wiener-Kolmogorov formula, expressing the optimal prediction in terms of "innovations," does obtain.

Since $|z_j| > \sqrt{\beta}$ for $j = 1, \ldots, m$ it follows that $|\lambda_j| < 1/\sqrt{\beta}$ for $j = 1, \ldots, m$. Using (1.11), we can express the factorization (1.10) as

$$[h + d(\beta z^{-1})d(z)] = c_0^2(1 - \lambda_1 z)\cdots(1 - \lambda_m z)(1 - \lambda_1 \beta z^{-1})\cdots(1 - \lambda_m \beta z^{-1}).$$

In sum, we have constructed a factorization (1.10) of the characteristic polynominal for the Euler equation in which the zeros of $c(z)$ exceed $\beta^{-1/2}$ in modulus, and the zeros of $c(\beta z^{-1})$ are less than $\beta^{-1/2}$ in modulus. Using (1.10), we now write the Euler equation as

$$c(\beta L^{-1})\, c(L)\, y_t = a_t. \tag{1.12}$$

The unique solution of the Euler equation that satisfies condition (1.6) is given by

$$c(L)\, y_t = c(\beta L^{-1})^{-1} a_t. \tag{1.13}$$

This can be established by using an argument paralleling that in Sargent [1987, chapter IX]. To exhibit the solution in a form paralleling that of Sargent [1987], we use (1.11) to write (1.13) as

$$(1 - \lambda_1 L)\cdots(1 - \lambda_m L)y_t = \frac{c_0^{-2} a_t}{(1 - \beta\lambda_1 L^{-1})\cdots(1 - \beta\lambda_m L^{-1})}. \tag{1.14}$$

Using partial fractions,[4] we can write the characteristic polynomial on the right side of (1.14) as

$$\frac{c_0^{-2}}{(1 - \lambda_1\beta L^{-1})\cdots(1 - \lambda_m\beta L^{-1})} = \sum_{j=1}^{m} \frac{A_j}{1 - \lambda_j \beta L^{-1}}$$

where

$$A_j = \frac{c_0^{-2}}{\prod_{i \neq j}(1 - \frac{\lambda_i}{\lambda_j})}.$$

Then (1.14) can be written

$$(1 - \lambda_1 L)\cdots(1 - \lambda_m L)y_t = \sum_{j=1}^{m} \frac{A_j}{1 - \lambda_j \beta L^{-1}} a_t$$

or

$$(1 - \lambda_1 L)\cdots(1 - \lambda_m L)y_t = \sum_{k=0}^{m} (\beta\lambda_j)^k\, a_{t+k}. \tag{1.15}$$

[4] See Sargent [1987] or Gabel and Roberts [1973]

Equation (1.15) expresses the optimum sequence for y_t in terms of m lagged y's, and m weighted infinite geometric sums of future a_t's. Furthermore, equation (1.15) is the unique solution of the Euler equation that satisfies the initial conditions and condition (1.6). In effect, condition (1.6) compels us to solve the "unstable" roots of $[h + d(\beta z^{-1})d(z)]$ forward (see Sargent [ ]). The step of factoring the polynomial $[h + d(\beta z^{-1})\, d(z)]$ into $c(\beta z^{-1})c(z)$, where the zeros of $c(z)$ are outside the unit circle, is central to solving the problem.

We note two features of the solution (1.15). First, since $|\lambda_j| < 1/\sqrt{\beta}$ for all j, it follows that $(\lambda_j\ \beta) < \sqrt{\beta}$. Therefore, the assumption that $\{a_t\}$ is of exponential order less than $1/\sqrt{\beta}$ is sufficient to guarantee that the geometric sums of future a_t's on right side of (1.15) converge. We immediately see that those sums will converge under the weaker condition that $\{a_t\}$ is of exponential order less than ϕ^{-1} where $\phi = \max\{\beta\lambda_i, i = 1, \dots, m\}$.

Second, note that with a_t identically zero, (1.15) implies that in general $|y_t|$ eventually grows exponentially at a rate given by $\max_i |\lambda_i|$. The condition $\max_i |\lambda_i| < 1/\sqrt{\beta}$ guarantees that condition (1.6) is satisfied. In fact, $\max_i |\lambda_i| < 1/\sqrt{\beta}$ is a necessary condition for (1.6) to hold. Were (1.6) not satisfied, the objective function diverges to $-\infty$, implying that the y_t path could not be optimal. For example, with $a_t = 0$, for all t, it is easy to describe a naive (nonoptimal) policy for $\{y_t, t > 0\}$ that gives a finite value of (1). We can simply let $y_t = 0$ for $t > 0$. This policy involves at most m nonzero values of hy_t^2 and $[d(L)y_t]^2$, and so yields a finite value of (1.1). Therefore it is easy to dominate a path that violates (1.6).

3. Undiscounted Problems

It is worthwhile focusing on a special case of the problem of Section 2, the undiscounted problem that emerges when $\beta = 1$. In this case, the Euler equation is

$$\left(h + d(L^{-1})d(L)\right) y_t = a_t.$$

The factorization of the characteristic polynomial (1.10) becomes

$$\left(h + d(z^{-1})d(z)\right) = c(z^{-1})\, c(z) \tag{1.19}$$

where

$$c(z) = c_0(1 - \lambda_1 z) \dots (1 - \lambda_m z)$$

$$c_0 = \left[(-1)^m z_0 z_1 \dots z_m\right]$$

$$|\lambda_j| < 1 \text{ for } j = 1, \dots, m.$$

The solution of the problem becomes

$$(1-\lambda_1 L)\cdots(1-\lambda_m L)y_t = \sum_{j=1}^{m} A_j \sum_{K=0}^{\infty} \lambda_j^k a_{t+k}.$$

Discounted problems can always be converted into undiscounted problems via a simple transformation. Thus consider problem (1) with $0 < \beta < 1$. Define the transformed variables

$$\tilde{a}_t = \beta^{t/2} a_t, \ \tilde{y}_t = \beta^{t/2} y_t. \tag{1.20}$$

Then notice that $\beta^t [d(L)y_t]^2 = [\tilde{d}(L)\tilde{y}_t]^2$ with $\tilde{d}(L) = \sum_{j=0}^{m} \tilde{d}_j L^j$ and $\tilde{d}_j = \beta^{j/2} d_j$. Then the original criterion function (1.1) is equivalent with

$$\lim_{N\to\infty} \sum_{t=0}^{N} \{\tilde{a}_t \tilde{y}_t - \frac{1}{2} h \tilde{y}_t^2 - \frac{1}{2}[\tilde{d}(L)\tilde{y}_t]^2\} \tag{1.1'}$$

which is to be maximized over sequences $\{\tilde{y}_t,\ t = 0, \ldots\}$ subject to $\tilde{y}_{-1}, \cdots, \tilde{y}_{-m}$ given and $\{\tilde{a}_t,\ t = 1, \ldots\}$ a known bounded sequence.

The Euler equation for this problem is $[h + \tilde{d}(L^{-1})\tilde{d}(L)]\tilde{y}_t = \tilde{a}_t$. The solution of this problem is

$$(1-\tilde{\lambda}_1 L)\cdots(1-\tilde{\lambda}_m L)\tilde{y}_t = \sum_{j=1}^{m} \tilde{A}_j \sum_{k=0}^{\infty} \tilde{\lambda}_j^k \tilde{a}_{t+k}$$

or

$$\tilde{y}_t = \tilde{f}_1 \tilde{y}_{t-1} + \cdots + \tilde{f}_m \tilde{y}_{t-m} + \sum_{j=1}^{m} \tilde{A}_j \sum_{k=0}^{\infty} \tilde{\lambda}_j^k \tilde{a}_{t+k} \tag{1.21}$$

where $\tilde{c}(z^{-1})\tilde{c}(z) = h + \tilde{d}(z^{-1})\tilde{d}(z)$, and where

$$[(-1)^m \tilde{z}_0 \tilde{z}_1 \ldots \tilde{z}_m]^{1/2}(1-\tilde{\lambda}_1 z)\ldots(1-\tilde{\lambda}_m z) = \tilde{c}(z), \quad \text{where } |\tilde{\lambda}_j| < 1.$$

We leave it to the reader to show that (1.21) implies the equivalent form of the solution

$$y_t = f_1 y_{t-1} + \cdots + f_m y_{t-m} + \sum_{j=1}^{m} A_j \sum_{k=0}^{\infty} (\lambda_j \beta)^k a_{t+k} \tag{1.22}$$

where

$$f_j = \tilde{f}_j \beta^{-j/2}, \ A_j = \tilde{A}_j, \ \lambda_j = \tilde{\lambda}_j \beta^{-1/2}. \tag{1.23}$$

By making use of the transformations (1.20) and the inverse formulas (1.23), it is always possible to solve a discounted problem by first solving a related undiscounted problem.

4. Infinite Dimensional Prediction and Signal Extraction

We now consider two related prediction and filtering problems. We let Y_t be a univariate m^{th} order moving average, covariance stationary stochastic process,

$$Y_t = d(L)u_t \tag{1.24}$$

where $d(L) = \sum_{j=0}^{m} d_j L^j$, and u_t is a serially uncorrelated stationary random process satisfying

$$\begin{aligned} &Eu_t = 0 \\ &Eu_t u_s = \begin{cases} 1 & t = s \\ 0 & t \neq s \end{cases} \end{aligned} \tag{1.25}$$

We impose no conditions on the zeros of $d(z)$. A second covariance stationary process is X_t given by

$$X_t = Y_t + \varepsilon_t \tag{1.26}$$

where ε_t is a serially uncorrelated stationary random process with $E\varepsilon_t = 0$ and

$$E\varepsilon_t\varepsilon_s = \begin{cases} h > 0 & t = s \\ 0 & t \neq s \end{cases}$$

It is assumed that $E\varepsilon_t u_s = 0$ for all t and s.

The *linear least squares prediction problem* is to find the random variable $\widehat{X}_{t+j}$ among linear combinations of $\{X_t,\ X_{t-1},\ldots\}$ that minimizes $E\{\widehat{X}_{t+j}\}^2$. That is, the problem is to find a $\gamma_j(L) = \sum_{k=0}^{\infty} \gamma_{jk} L^k$ such that $\sum_{k=0}^{\infty} |\gamma_{jk}|^2 < \infty$ and such that $E\{\gamma_j(L)X_t - X_{t+j}\}^2$ is minimized.

The *linear least squares filtering problem* is to find a $b(L) = \sum_{j=0}^{\infty} b_j L^j$ such that $\sum_{j=0}^{\infty} |b_j|^2 < \infty$ and such that $E\{b(L)X_t - Y_t\}^2$ is minimized. Interesting versions of these problems related to the permanent income theory were studied by Muth [ ].

These problems are solved as follows. The covariograms of Y and X and their cross covariogram are, respectively, defined as

$$\begin{aligned} C_X(\tau) &= EX_tX_{t-\tau} \\ C_Y(\tau) &= EY_tY_{t-\tau} \qquad \tau = \pm 1, \pm 2, \ldots \\ C_{Y,X}(\tau) &= EY_tX_{t-\tau} \end{aligned} \tag{1.27}$$

The covariance and cross covariance generations functions are defined as

$$
\begin{aligned}
g_X(z) &= \sum_{\tau=-\infty}^{\infty} C_X(\tau) z^{\tau} \\
g_Y(z) &= \sum_{\tau=-\infty}^{\infty} C_Y(\tau) z^{\tau} \qquad (1.28)\\
g_{YX}(z) &= \sum_{\tau=-\infty}^{\infty} C_{YX}(\tau) z^{\tau}.
\end{aligned}
$$

The generating functions can be computed by using the following facts. Let v_{1t} and v_{2t} be two mutually and serially uncorrelated white noises with unit variances, i.e., $Ev_{1t}^2 = Ev_{2t}^2 = 1, Ev_{1t} = Ev_{2t} = 0, Ev_{1t}v_{2s} = 0$ for all t and s, $Ev_{1t}v_{1t-j} = Ev_{2t}v_{2t-j} = 0$ for all $j \neq 0$. Let x_t and y_t be two random process given by

$$
\begin{aligned}
y_t &= A(L)v_{1t} + B(L)v_{2t} \\
x_t &= C(L)v_{1t} + D(L)v_{2t}.
\end{aligned}
$$

Then, as shown for example in Sargent [, ch. 11], we have

$$
\begin{aligned}
g_y(z) &= A(z)A(z^{-1}) + B(z)B(z^{-1}) \\
g_x(z) &= C(z)C(z^{-1}) + D(z)D(z^{-1}) \qquad (1.29)\\
g_{yx}(z) &= A(z)C(z^{-1}) + B(z)D(z^{-1}).
\end{aligned}
$$

Applying these formulas to (1.24)-(1.27), we have

$$
\begin{aligned}
g_Y(z) &= d(z)d(z^{-1}) \\
g_X(z) &= d(z)d(z^{-1}) + h \qquad (1.30)\\
g_{YX}(z) &= d(z)d(z^{-1}).
\end{aligned}
$$

The key step in obtaining solutions to our problems is to factor the covariance generating function of X, $g_X(z)$. The solutions of our problems are given by formulas due to Wiener and Kolmogorov. These formulas utilize the Wold moving average representation of the X_t process,[5]

$$X_t = c(L)\eta_t \qquad (1.31)$$

[5] The existence of which is assured by Wold's representation theorem. See, for example, Sargent [, ch. XI].

where $c(L) = \sum_{j=0}^{m} c_j L^j$, where

$$c_0 \eta_t = X_t - \widehat{E}[X_t | X_{t-1}, X_{t-2}, \ldots], \tag{1.32}$$

where $\widehat{E}$ is the linear least squares projection operator. Equation (1.32) is the condition that $c_0 \eta_t$ can be the one-step ahead error in predicting X_t from its own past values. Condition (1.32) requires that η_t lie in the closed linear space spanned by $[X_t, X_{t-1}, \ldots]$. This will be true if and only if the zeros of $c(z)$ do not lie inside the unit circle. It is an implication of (1.32) that η_t is a serially uncorrelated random process, and that a normalization can be imposed so that $E\eta_t^2 = 1$. Consequently, an implication of (1.31) is that the covariance generating function of X_t can be expressed as

$$g_X(z) = c(z) c(z^{-1}) \tag{1.33}$$

It remains to discuss how $c(L)$ is to be computed. Combining (1.29) and (1.33) gives

$$d(z) d(z^{-1}) + h = c(z) c(z^{-1}). \tag{1.34}$$

Now equation (1.34) is identical with (1.10). Further, the conditions that (1.31) imposes on $c(z)$, that its zeros not lie inside the unit circle, are identical with those imposed in (1.9). Therefore, we have already showed constructively how to factor the covariance generating function $g_X(z) = d(z) d(z^{-1}) + h$.

We now introduce the "annihilation operator:"

$$\left[\sum_{j=-\infty}^{\infty} f_j L^j \right]_+ \equiv \sum_{j=0}^{\infty} f_j L^j. \tag{1.35}$$

In words, $[\quad]_+$ means "ignore negative powers of L." We have defined the solution of the prediction problem as $\widehat{E}[X_{t+j} | X_t, X_{t-1}, \ldots] = \gamma_j(L) X_t$. Assuming that the roots of $c(z) = 0$ all lie outside the unit circle, the Wiener-Kolmogorov formula for $\gamma_j(L)$ holds:

$$\gamma_j(L) = \left[\frac{c(L)}{L^j} \right]_+ c(L)^{-1}. \tag{1.36}$$

We have defined the solution of the filtering problem as $\hat{E}[Y_t \mid X_t, X_{t-1}, \ldots] = b(L) X_t$. The Wiener-Kolomogorov formula for $b(L)$ is

$$b(L) = \left(\frac{g_{YX}(L)}{c(L^{-1})} \right)_+ c(L)^{-1} \tag{1.37}$$

or

$$b(L) = \left(\frac{d(L)d(L^{-1})}{c(L^{-1})} \right)_+ c(L)^{-1}.$$

Formulas (1.36) and (1.37) are discussed in detail in Whittle [ ] and Sargent [ ]. The interested reader can there find several examples of the use of these formulas in economics. Some classic examples using these formulas are due to Muth [ ].

As an example of the usefulness of formula (1.37), we let X_t be a stochastic process with Wold moving average representation

$$X_t = c(L)\eta_t$$

where $E\eta_t^2 = 1$, and $c_0\eta_t = X_t - \widehat{E}[X_t|X_{t-1},\ldots], c(L) = \sum_{j=0}^{m} c_j L$. Suppose that at time t, we wish to predict a geometric sum of future X's, namely

$$y_t \equiv \sum_{j=0}^{\infty} \delta^j X_{t+j} = \frac{1}{1-\delta L^{-1}} X_t$$

given knowledge of $X_t, X_{t-1}, \ldots$. We shall use (37) to obtain the answer. Using the standard formulas (1.29), we have that

$$g_{yx}(z) = (1 - \lambda z^{-1})c(z)c(z^{-1})$$

$$g_x(z) = c(z)c(z^{-1}).$$

Then, (1.37) becomes

(1.38)
$$b(L) = \left[\frac{c(L)}{1-\delta L^{-1}}\right]_+ c(L)^{-1}.$$

In order to evaluate the term in the annihilation operator, we use the following result from Hansen and Sargent [ ].

Proposition: Let $g(z) = \sum_{j=0}^{\infty} g_j z^j$ where $\sum_{j=0}^{\infty} |g_j|^2 < +\infty$. Let $h(z^{-1}) = (1 - \delta_1 z^{-1})\ldots(1-\delta_n z^{-1})$, where $|\delta_j| < 1$, for $j = 1, \ldots, n$. Then

(1.39)
$$\left[\frac{g(z)}{h(z^{-1})}\right]_+ = \frac{g(z)}{h(z^{-1})} - \sum_{j=1}^{n} \frac{\delta_j^n g(\delta_j)}{\prod_{\substack{k=1 \\ k\neq j}}^{n} (\delta_k - \delta_j)} \left(\frac{1}{z-\delta_j}\right)$$

and, alternatively,

(1.40)
$$\left[\frac{g(z)}{h(z^{-1})}\right]_+ = \sum_{j=1}^{n} B_j \left(\frac{zg(z) - \delta_j g(\delta_j)}{z - \delta_j}\right)$$

where $B_j = 1/\prod_{k\neq j}^{n}(1 - \delta_k \delta_j)$.

Applying formula (1.40) of the proposition to evaluating (1.38) with $g(z) = c(z)$ and $h(z^{-1}) = 1 - \delta z^{-1}$ gives

$$b(L) = \left[\frac{Lc(L) - \delta c(\delta)}{L - \delta}\right]c(L)^{-1}$$

or

$$b(L) = \left[\frac{1 - \delta c(\delta)L^{-1}c(L)^{-1}}{1 - \delta L^{-1}}\right].$$

Thus, we have

$$\hat{E}\Big[\sum_{j=0}^{\infty} \delta^j x_{t+j} | x_t, x_{t-1} \ldots\Big] = \left[\frac{1 - \delta c(\delta)L^{-1}c(L)^{-1}}{1 - \delta L^{-1}}\right]x_t. \tag{1.41}$$

This formula is useful in solving stochastic versions of problem (1.1) in which the randomness emerges because a_t is a stochastic process. The problem is to maximize

$$E_0 \lim_{N\to\infty} \sum_{t-0}^{N} \beta^t\Big[a_t Y_t - \frac{1}{2}hy_t^2 - \frac{1}{2}[d(L)y_t]^2\Big] \tag{1.42}$$

where E_t is mathematical expectation conditioned on information known at t, and where $\{a_t\}$ is a covariance stationary stochastic process with Wold moving average representation

$$a_t = \bar{c}(L)\,\eta_t$$

where

$$\bar{c}(L) = \sum_{j=0}^{\bar{n}} \bar{c}_j\, L^j,$$

and $\eta_t = a_t - \hat{E}[a_t | a_{t-1}, \ldots]$.

The problem is to maximize (1.42) with respect to a contingency plan expressing y_t as a function of information known at t, which is assumed to be $(y_{t-1},\ y_{t-2}, \ldots, a_t,\ a_{t-1}, \ldots)$. The solution of this problem can be achieved in two steps. First, ignoring the uncertainty, we can solve the problem assuming that a_t is a known sequence. The solution is, from above,

$$c(L)y_t = c(L^{-1})^{-1}a_t$$

or

$$(1 - \lambda_1 L)\ldots(1 - \lambda_m L)y_t = \sum_{j=1}^{m} A_j \sum_{k=0}^{\infty}(\lambda\beta)^k a_{t+k}. \tag{1.43}$$

Second, the solution of the problem under uncertainty is obtained by replacing the terms on the right-hand side of the above expressions with their linear least squares predictors. Using (1.41) and (1.43), we have the following solution

$$(1-\lambda_1 L)\dots(1-\lambda_m L)y_t = \sum_{j-1}^{m} A_j\Big[\frac{1-\beta\lambda_j \bar{c}(\beta\lambda_j)L^{-1}\bar{c}(L)^{-1}}{1-\beta\lambda_j L^{-1}}\Big]a_t.$$

5. Finite Dimensional Control

We briefly study the finite horizon version of our optimization problem, using matrix methods. For simplicity, we shall focus on the special case in which $m = 1$, although it should be clear how things will generalize to the case in which $m > 1$.[6] We want to solve the system of $N+1$ linear equations.

$$\begin{aligned} &[h + d(\beta L^{-1})\,d(L)]y_t = a_t, \quad t = 0,\ 1,\ \dots,\ N-1 \\ &\beta^N[a_N - h\,y_n - d_0\,d(L)y_N] = 0 \end{aligned} \tag{1.44}$$

where $d(L) = d_0 + d_1 L$. These equations are to be solved for $y_0, y_1, \dots, y_{n-1}$ and y_N as functions of $a_0, a_1, \dots, a_{N-1}$ and a_N. Let $\phi(L) = \phi_0 + \phi_1 L + \beta\phi_1 L^{-1} = h + d(\beta L^{-1})d(L) = (h + d_0^2 + d_1^2) + d_1 d_0 L + d_1 d_0 \beta L^{-1}$. Then we can represent (1.44) as the matrix equation

$$\begin{bmatrix} (\phi_0 - d_1^2) & \phi_1 & 0 & 0 & \dots & \dots & 0 \\ \beta\phi_1 & \phi_0 & \phi_1 & 0 & \dots & \dots & 0 \\ 0 & \beta\phi_1 & \phi_0 & \phi_1 & \dots & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \dots & \dots & \dots & \beta\phi_1 & \phi_0 & \phi_1 \\ 0 & \dots & \dots & \dots & 0 & \beta\phi_1 & \phi_0 \end{bmatrix} \begin{bmatrix} y_N \\ y_{N-1} \\ y_{N-2} \\ \vdots \\ y_1 \\ y_0 \end{bmatrix} = \begin{bmatrix} a_N \\ a_{N-1} \\ a_{N-2} \\ \vdots \\ a_1 \\ a_0 - \phi_1 y_{-1} \end{bmatrix} \tag{1.45}$$

or

$$W\bar{y} = \bar{a}. \tag{1.46}$$

Notice how we have chosen to arrange the y_t's in reverse time order. The matrix W on the left side of (1.45) is "almost" a toeplitz matrix, there being two sources of deviation from the toeplitz form. First, the (2,1) element differs from the remaining diagonal elements, reflecting the terminal condition. Second, the subdiagonal elements equal β time the superdiagonal elements.

[6] See exercise number .

The solution of (1.46) can be expressed in the form

$$\bar{y} = W^{-1}\bar{a}, \tag{1.47}$$

which represents each element y_t of $\bar{y}$ of a function of the entire vector $\bar{a}$. That is, y_t is a function of past, present, and future values of a's, as well as of the initial condition y_{-1}.

An alternative way to express the solution to (1.45) or (1.46) is in so called feedback—feedforward form. The idea here is to find a solution expressing y_t as a function of *past* y's and *current* and *future* a_s's. To achieve this solution, one can use an "LU" decomposition of W. There always exists a decomposition of W of the form

$$W = LU \tag{1.48}$$

where L is an $(N+1)\times(N+1)$ lower triangular matrix, and U is an $(N+1)\times(N+1)$ upper trangular matrix. The factorization can be normalized so that the diagonal elements of U are unity. Using representation (1.48) in equation (1.47) we obtain

$$U\bar{y} = L^{-1}\bar{a}. \tag{1.49}$$

Since L^{-1} is lower trangular, this representation expresses y_t as a function of lagged y's (via the term $U\bar{y}$) and current and future a's (via the term $L^{-1}\bar{a}$). Because there are zeros everwhere in the matrix on the left of (1.45) except in the diagonal, superdiagonal, and subdiagional, the LU decomposition takes L to be zero except in the diagional and the leading subdiagonal, while U is zero except on the diagonal and the superdiagional. Thus, (1.49) has the form

$$\begin{bmatrix} 1 & U_{12} & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & U_{23} & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & U_{34} & \dots & 0 & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 & U_{N-1,N} \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 \end{bmatrix} \begin{bmatrix} y_N \\ y_{N-1} \\ y_{N-2} \\ y_{N-3} \\ \vdots \\ y_1 \\ y_0 \end{bmatrix} =$$

$$\begin{bmatrix} L_{11}^{-1} & 0 & 0 & \dots & 0 \\ L_{21}^{-1} & L_{22}^{-1} & 0 & \dots & 0 \\ L_{31}^{-1} & L_{32}^{-1} & L_{33}^{-1} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ L_{N,1}^{-1} & L_{N,2}^{-1} & L_{N,3}^{-1} & \dots & 0 \\ L_{N+1,1}^{-1} & L_{N+1,2}^{-1} & L_{N+1,3}^{-1} & \dots & L_{N+1\,N+1}^{-1} \end{bmatrix} \begin{bmatrix} a_N \\ a_{N-1} \\ a_{N-2} \\ \vdots \\ a_1 \\ a_0 - \phi_1 y_{-1} \end{bmatrix} \tag{1.50}$$

where L_{ij}^{-1} is the (i,j) element of L^{-1} and U_{ij} is the (i,j) element of U.

We briefly indicate how this approach extends to the problem with $m > 1$. Assume that $\beta = 1$. Let D_{m+1} be the $(m+1) \times (m+1)$ symmetric matrix whose elements are determined from the following formula:

$$D_{jk} = d_0 d_{k-j} + d_1 d_{k-j+1} + \ldots + d_{j-1} d_{k-1}, \qquad k \geq j.$$

Let I_{m+1} be the $(m+1) \times (m+1)$ identity matrix. Let ϕ_j be the coefficients in the expansion $\phi(L) = h + d(L^{-1})d(L)$. Then the first order conditions (1.4) and (1.5) can be expressed as:

$$(D_{m+1} + hI_{m+1}) \begin{bmatrix} y_N \\ y_{N-1} \\ \vdots \\ y_{N-m} \end{bmatrix} = \begin{bmatrix} a_N \\ a_{N-1} \\ \vdots \\ a_{N-m} \end{bmatrix}$$

$$\phi_m y_N + \phi_{m-1} y_{N-1} + \ldots + \phi_0 y_{N-m} + \phi_1 y_{N-m-1} + \ldots + \phi_m y_{N-2m} = a_{N-m-1}$$

$$\phi_m y_{N-1} + \phi_{m-1} y_{N-2} + \ldots + \phi_0 y_{N-m-1} + \phi_1 y_{N-m-2} + \ldots + \phi_m y_{N-2m-1} = a_{N-m-2}$$

$$\vdots$$

$$\phi_m y_{m+1} + \phi_{m-1} y_m + + \ldots + \phi_0 y_1 + \phi_1 y_0 + \phi_m y_{-m+1} = a_1$$

$$\phi_m y_m + \phi_{m-1} y_{m-1} + \phi_{m-2} + \ldots + \phi_0 y_0 + \phi_1 y_{-1} + \ldots + \phi_m y_{-m} = a_0$$

The matrix on the left of this equation is "almost" toeplitz, the exception being the leading $m \times m$ sub matrix in the upper left hand corner. As before, we can express equation as

$$W\bar{y} = \bar{a}. \tag{1.51}$$

We can represent the solution in feedback-feedforward form by obtaining a decomposition LU $= W$, and obtain

$$U\bar{y} = L^{-1}\bar{a}.$$

$$\sum_{j=0}^{t} U_{t,t-j}\; y_{t-j} = \sum_{j=0}^{N-t} L_{t,t+j}^{-1}\; a_{t+j} \tag{1.52}$$

where $L_{t,s}^{-1}$ is the element in the $t+m+1$ row and $s+m+1$ column of L, with a similar convention of $U_{t,s}$.

The left side of equation (1.52) is the "feedback" part of the optimal control law for y_t, while the right-hand side is the "feedforward" part. We note that there is a different control law for each t. Thus, in the finite horizon case, the optimal control law is time dependent.

It is natural to suspect that as $N \to \infty$, (1.52) becomes equivalent to the solution of our infinite horizon problem, which we have expressed as

$$c(L)y_t = c(\beta L^{-1})^{-1} a_t,$$

so that as $N \to \infty$ we expect that for each fixed $t, L^{-1}_{t,t-j} \to c_j$ and $U_{t,t+j}$ approaches the coefficient on L^{-j} in the expansion of $c(\beta L^{-1})$. This suspicion is true under general conditions which we shall study later. For now, we note that by creating the matrix W for large N and factoring it into the LU form, good approximations to $c(L)$ and $c(\beta L^{-1})^{-1}$ can be obtained.

6. Finite Dimensional Prediction

Let $(x_1, x_2, \ldots x_T)' = x$ be a $T \times 1$ vector of random variables with mean $Ex = 0$ and covariance matrix $Exx' = V$. Here V is a $T \times T$ positive definite matrix. We shall regard the random variables as being ordered in time, so that x_t is thought of as the value of some economic variable at time t. For example, x_t could be generated by the random process described in Section 5. In this case, V_{ij} is given by the coefficient on $z^{|i-j|}$ in the expansion of $g_x(z) = d(z)\, d(z^{-1}) + h$, which equals $h + \sum_{k=0}^{\infty} d_k d_{k+|i-j|}$. We shall be interested in constructing $j-$step ahead linear least squares predictors of the form

$$\widehat{E}[x_{T-j}, x_{T-j+1}, \ldots, x_1]$$

where $\widehat{E}$ is the linear least squares projection operator.

The solution of this problem is clearly exhibited by first constructing an orthonormal basis of random variables ε for x. Since V is a positive definite and symmetric, we know that there exists a (Cholesky) decomposition of V such that

$$V = L^{-1}(L^{-1})'$$

or

$$L\, V\, L' = I$$

where L is lower-trangular, and therefore so is L^{-1}. Form the random variable $Lx = \varepsilon$. Then ε is an orthonormal basis for x, since L is nonsingular, and $E\varepsilon\varepsilon' = LExx'L' = I$.

It is convenient to write out the equations $Lx = \varepsilon$ and $L^{-1}\varepsilon = x$.

$$
\begin{aligned}
L_{11}x_1 &= \varepsilon_1 \\
L_{21}x_1 + L_{22}x_2 &= \varepsilon_2 \\
&\vdots \\
L_{T1}x_1 \ldots + L_{TT}x_T &= \varepsilon_T
\end{aligned}
\tag{1.53}
$$

or

$$\sum_{j=0}^{t-1} L_{t,t-j}\, x_{t-j} = \varepsilon_t, \quad t = 1, 2, \ldots T \tag{1.54}$$

We also have

$$x_t = \sum_{j=0}^{t-1} L^{-1}_{t,t-j}\, \varepsilon_{t-j}. \tag{1.55}$$

Notice from (1.55) that x_t is in the space spanned by $\varepsilon_t, \varepsilon_{t-1}, \ldots, \varepsilon_1$, and from (1.54) that ε_t is in the space spanned by $x_t, x_{t-1}, \ldots, x_1$.

Therefore, we have that for $t-1 \geq m \geq 1$

$$\widehat{E}\,[x_t \mid x_{t-m}, x_{t-m-1}, \ldots, x_1] = \widehat{E}[x_t \mid \varepsilon_{t-m}, \varepsilon_{t-m+1}, \ldots, \varepsilon_1]. \tag{1.56}$$

For $t-1 \geq m \geq 1$ rewrite (1.55) as

$$x_t = \sum_{j=0}^{m-1} L^{-1}_{t,t-j}\, \varepsilon_{t-j} + \sum_{j=m}^{t-1} L^{-1}_{t,t-j}\, \varepsilon_{t-j}. \tag{1.57}$$

Representation (1.50) is an orthogonal decomposition of x_t into a part $\sum_{j=m}^{t-1} L^{-1}_{t,t-j}\, \varepsilon_{t-j}$ that lies in the space spanned by $[x_{t-m}, x_{t-m+1}, \ldots, x_1]$, and an orthogonal component not in this space. It immediately follows from the "orthogonality principle" of least squares (see Papoulis [ ] or Sargent [ ]) that

$$
\begin{aligned}
\widehat{E}[x_t \mid x_{t-m}, x_{t-m+1}, \ldots x_1] &= \sum_{j=m}^{t-1} L^{-1}_{t,t-j}\, \varepsilon_{t-j} \\
&= [L^{-1}_{t,1}\, L^{-1}_{t,2}, \ldots, L^{-1}_{t,t-m}\; 0\; 0 \ldots 0]L\,x.
\end{aligned}
\tag{1.56}
$$

This can be interpreted as a finite-dimensional version of the Wiener-Kolmogorov m-step ahead predict on formula.

We can use (1.51) to represent the linear least squares projection of the vector x conditioned on the first s observations $[x_s, x_{s-1} \ldots, x_1]$. We have

$$\hat{E}[x \mid x_s, x_{s-1}, \ldots, x_1] = L^{-1} \begin{bmatrix} I_s & 0 \\ 0 & 0_{(t-s)} \end{bmatrix} Lx.$$

This formula will be convenient in representing the solution of control problems under uncertainty.

Equation (1.55) can be recognized as a finite dimensional version of a moving average representation. Equation (1.54) can be viewed as a finite dimension version of an autoregressive representation. Notice that even if the x_t process is covariance stationary, so that V is such that V_{ij} depends only on $|i-j|$, the coefficients in the moving average representation are time-dependent, there being a different moving average for each t. If x_t is a covariance stationary process, the last row of L^{-1} converges to the coefficients in the Wold moving average representation for $\{x_t\}$ as $T \to \infty$. Further, if x_t is covariance stationary, for fixed k and $j > 0$, $L^{-1}_{T,T-j}$ converges to $L^{-1}_{T-k,T-k-j}$ as $T \to \infty$. That is, the "bottom" rows of L^{-1} converge to each other and to the Wold moving average coefficients as $T \to \infty$.

This last observation gives one simple and widely-used practical way of forming a finite T approximation to a Wold moving average representation. First, form the covariance matrix $Exx' = V$, then obtain the Cholesky decomposition $L^{-1}L^{-1'}$ of V, which can be accomplished quickly on a computer. The last row of L^{-1} gives the approximate Wold moving average coefficients. This method can readily be generalized to multivariate systems.

7. Combined Finite Dimensional Control and Prediction

Consider the finite-dimensional control problem, maximize

$$E \sum_{t=0}^{N} \{a_t y_t - \frac{1}{2} h y_t^2 - \frac{1}{2}[d(L) y_t]^2\}, \; h > 0$$

where $d(L) = d_0 + d_1 L + \ldots + d_m L^m$, L is the lag operator, $\bar{a} = [a_N, a_{N-1} \ldots, a_1, a_0]'$ a random vector with mean zero and $E\, a_t a_t' = V$. The variables $y_{-1}, \ldots y_{-m}$ are given. The maximization is over choices of $y_0, y_1 \ldots, y_N$, where y_t is required to be a linear function of $\{y_{t-s-1};\ a_{t-s};\ t-1 \geq s \geq 0\}$.

We saw in section 5 that the solution of this problem under certainty could be represented in feedback-feedforward form

$$U\bar{y} = L\bar{a}.$$

Using a version of formula 1.58, we can express $\hat{E}[\bar{a} \mid a_s, a_{s-1}, \ldots, a_0]$ as

$$\hat{E}[\bar{a} \mid a_s, a_{s-1}, \ldots, a_0] = \tilde{U}^{-1} \begin{bmatrix} 0 & 0 \\ 0 & I_{(s+1)} \end{bmatrix} \tilde{U}\bar{a}$$

where $I_{(s+1)}$ is the $(s+1) \times (s+1)$ identity matrix, and $V = \tilde{U}^{-1}\tilde{U}^{-1'}$, where $\tilde{U}$ is the *upper* triangular Cholesky factor of the covariance matrix V. (We have reversed the time axis in dating the a's relative to section 5. The time axis can be reversed in representation ( ) by replacing L with L^T.)

The optimal decision rule to use at time $0 \leq t \leq N$ is then given by the $(N - t + 1)^{\text{th}}$ row of

$$U\bar{y} = L^{-1}\tilde{U}^{-1} \begin{bmatrix} 0 & 0 \\ 0 & I_{(t+1)} \end{bmatrix} \tilde{U}\bar{a}.$$

Exercises

1. Consider solving a discounted version ($\beta < 1$) of problem (1.1), as follows. Use the transformations in footnote 2, to convert (1.1) to the undiscounted problem (1.1′) of footnote 2. Let the solution of (1.1′) in feedback form be

$$(1-\tilde{\lambda}_1 L)\cdots(1-\tilde{\lambda}_m L)\tilde{y}_t = \sum_{j=1}^{m}\tilde{A}_j\sum_{k=0}^{\infty}\tilde{\lambda}_j^k\tilde{a}_{t+k}$$

or

$$(*)\qquad \tilde{y}_t = \tilde{f}_1\tilde{y}_{t-1}+\cdots+\tilde{f}_m\tilde{y}_{t-m}+\sum_{j=1}^{m}\tilde{A}_j\sum_{k=0}^{\infty}\tilde{\lambda}_j^k\tilde{a}_{t+k}$$

where $h+\tilde{d}(z^{-1})\tilde{d}(z) = \tilde{c}(z^{-1})\tilde{c}(z)$ and $\tilde{c}(z) = [(-1)^m\tilde{z}_0\tilde{z}_1\cdots\tilde{z}_m]^{1/2}(1-\tilde{\lambda}_1 z)\cdots(1-\tilde{\lambda}_m z)$, where the $\tilde{z}_j$ are the zeros of $h+\tilde{d}(z^{-1})\tilde{d}(z)$. Prove that (*) implies that the solution for y_t in feedback form is

$$y_t = f_1 y_{t-1}+\ldots+f_m y_{t-m}+\sum_{j=1}^{m}A_j\sum_{k=0}^{\infty}\beta^k\lambda_j^k a_{t+k}$$

where $f_j = \tilde{f}_j\beta^{-j/2}$, $A_j = \tilde{A}_j$, and $\lambda_j = \tilde{\lambda}_j\beta^{-1/2}$.

2. Solve the optimal control problem, maximize

$$\sum_{t=0}^{2}\left\{a_t y_t - \frac{1}{2}[(1-2L)y_t]^2\right\}$$

subject to y_{-1} given, and $\{a_t\}$ a known bounded sequence. Express the solution in the "feedback form" (1.15), giving numerical values for the coefficients. Make sure that the boundary conditions (1.5) are satisfied. (Note: this problem differs from the problem in the text in one important way: instead of $h > 0$ in (1.1), $h = 0$. This has an important influence on the solution.)

3. Solve the infinite time optimal control problem to maximize

$$\lim_{N\to\infty}\sum_{t=0}^{N} -\frac{1}{2}[(1-2L)y_t]^2,$$

subject to y_{-1} given. Prove that the solution is

$$y_t = 2y_{t-1} = 2^{t+1}y_{-1}\qquad t > 0.$$

4. Solve the infinite time problem, to maximize

$$\lim_{N\to\infty} \sum_{t=0}^{N} (.0000001)\, y_t^2 - \frac{1}{2}[(1-2L)y_t]^2$$

subject to y_{-1} given. Prove that the solution $y_t = 2y_{t-1}$ to problem (1.3) violates condition (1.6), and so is not optimal. Prove that the optimal solution is approximately

$$y_t = \frac{1}{2}y_{t-1} = \left(\frac{1}{2}\right)^{t+1} y_{-1}, \qquad t > 0.$$

5. Consider a stochastic process with moving average representation

$$x_t = (1-2L)\varepsilon_t$$

where ε_t is a serially uncorrelated random process with mean zero and variance unity. Use the Wiener-Kolmogorov foumula (1.36) to compute the linear least squares forecasts $E[x_{t+j} \mid x_t, x_{t-1}, \ldots]$, for $j = 1, 2$.

Hint: Let $\pi(z) = \sum_{j=0}^{m} \pi_j z^j$. Let $z_1, \ldots, z_k$ be the zeros of $\pi(z)$ that are inside the unit circle, $k < m$. Then define

$$\theta(z) = \pi(z)\left(\frac{(z_1 z - 1)}{(z - z_1)}\right)\left(\frac{(z_2 z - 1)}{(z - z_2)}\right)\cdots\left(\frac{(z_k z - 1)}{(z - z_k)}\right).$$

The term multiplying $\pi(z)$ is termed a "Blaschke factor." Then it can be proved directly that

$$\theta(z^{-1})\theta(z) = \pi(z^{-1})\pi(z)$$

and that the zeros of $\theta(z)$ are not inside the unit circle.

6. Consider a stochastic process X_t with moving average representation

$$X_t = (1 - \sqrt{2}L + L^2)\varepsilon_t$$

where ε_t is a serially uncorrelated random process with mean zero and variance unity.

   a. Find a Wold moving average representation for x_t.

   b. Use the Wiener-Kolomogorov formula (27) to compute the linear least squares forecasts $\widehat{E}[X_{t+j} \mid X_{t-1}, \ldots]$ for $j = 1, 2, 3$.

   (The hint to the previous problem is again useful.)

7. Let $Y_t = (1 - 2L)u_t$ where u_t is a mean zero white noise with $Eu_t^2 = 1$. Let

$$X_t = Y_t + \varepsilon_t$$

where ε_t is a serially uncorrleated white noise with $E\varepsilon_t^2 = 9$, and $E\varepsilon_t u_s = 0$ for all t and s.

a. Find the Wold moving average representation for X_t.

b. Find a formula for the A_{1j}'s in

$$E\widehat{X}_{t+1} \mid X_t, X_{t-1}, \ldots = \sum_{j=0}^{\infty} A_{1j} X_{t-j}$$

c. Find a formula for the A_{2j}'s in

$$\hat{E} X_{t+2} \mid X_t, X_{t-1}, \ldots = \sum_{j=0}^{\infty} A_{2j} X_{t-j}.$$

8. (A multiple variable control problem)

Consider the problem, maximize

$$\lim_{N\to\infty} \sum_{t=0}^{N} \beta^t \left\{ A_t' Y_t - \frac{1}{2} Y_t' H Y_t - \frac{1}{2}[D(L)Y_t]^2 \right\},\ 0 < \beta < 1,$$

where Y_t is an $(n \times 1)$ vector, $\{A_t, t = 0, 1, \ldots\}$ an $n \times 1$ vector of known sequences of exponential order less than $\beta^{1/2}$, $D(L) = D_0 + D_1 L + \ldots + D_m L^m$ where the D_j are $n \times n$ matrices, and H is an $n \times n$ positive definite matrix. The maximization is subject to $Y_{-1}, \ldots, Y_{-m}$ given, and is over infinite sequences for $\{Y_t, t = 0, 1, \ldots\}$.

a. Prove that the Euler equations are

$$[H + D(\beta L^{-1})' D(L)] Y_t = A_t$$

b. Give a boundary condition that generalizes (1.6).

c. Prove that if $\bar{z}$ is a zero of $|H + D(\beta\, z^{-1})' D(z)|$ then so is $\beta \bar{z}^{-1}$.

To solve the Euler equations subject to the boundary conditions it is necessary to achieve the factorization

$$[H + D(\beta z^{-1})' D(z)] = C(\beta z^{-1})' C(z)$$

where the zeros of $|C(z)|$ exceed β in modulus, and those of $|C(\beta z^{-1})|$ are less than β in modulus. Hansen and Sargent [ ] describe methods for achieving this factorization. The solution of the control problem can then be represented

$$C(L)Y_t = C(\beta L^{-1})' A_t.$$

9. (Multivariable Prediction)

Let Y_t be an $(n \times 1)$ vector stochastic process with moving average representation

$$Y_t = D(L)U_t$$

where $D(L) = \sum_{j+0}^{m} D_j L^J$, D_j an $n \times n$ matrix, U_t an $(n \times 1)$ vector white noise with

$$EU_t = 0 \qquad \text{forall } t$$
$$EU_t U_s' = \begin{cases} I & t = s \\ 0 & t \neq s \end{cases}$$

Let ε_t be an $n \times 1$ vector white noise with $E\varepsilon_t = 0$ for all t, $E\varepsilon_t U_s' = 0$ for all t and s and

$$E\varepsilon_t \varepsilon_{t-s}' = \begin{cases} H & t = s \\ 0 & t \neq s \end{cases}$$

where H is a positive definite matrix. Define the covariograms as $C_X(\tau) = EX_t X_{t-\gamma}'$, $C_Y(\tau) = EY_t Y_{t-\tau}'$, $C_{YX}(\tau) = EY_t X_{t-\tau}'$. Then define the matrix covariance generating function, as in (1.20), only interpret all the objects in (1.20) as matrices.

a. Show that the covariance generating functions are given by

$$g_y(z) = D(z)D(z^{-1})'$$
$$g_X(z) = D(z)D(z^{-1})' + H$$
$$g_{YX}(z) = D(z)D(z^{-1})'$$

b. A factorization of $g_X(z)$ can be found (see Rozanov [ ] or Whittle [ ]) of the form

$$D(z)D(z^{-1})' + H = C(z)C(z^{-1})', \quad C(z) = \sum_{j=0}^{m} C_j z^j$$

where the zeros of $|C(z)|$ do not lie inside the unit circle. A vector Wold moving average representation of X_t is then

$$X_t = C(L)\eta_t$$

where η_t is an $(n \times 1)$ vector white noise that is "fundamental" for X_t. That is, $X_t - \widehat{E}[X_t \mid X_{t-1}, X_{t-2} \ldots] = C_0 \, \eta_t$.

c. The optimum predictor of X_{t+j}, is

$$\widehat{E}[X_{t+j} \mid X_t, X_{t-1}, \ldots] = \left(\frac{C(L)}{L^j}\right)_+ \eta_t.$$

If $C(L)$ is invertible, i.e., if the zeros of det $C(z)$ lie strictly outside the unit circle, then this formula can be written

$$\widehat{E} X_{t+j} \mid X_t, X_{t-1}, \ldots = \left(\frac{C(L)}{L^j}\right)_+ C(L)^{-1} \, \eta_t.$$

References

Muth, John F., "Optimal Properties of Exponentially Weighted Forecasts," *Journal of the American Statistical Association*, 1960, vol. 55, no. 290.

Hansen, Lars Peter, and Thomas J. Sargent, "Formulating and Estimating Dynamics Linear Rational Expectations Models," *Journal of Economic Dynamics and Control*, 1980, vol. 2.

Sargent, Thomas J., *Macroeconomic Theory*, New York, Academic Press, 1979.

Whittle, Peter, *Regulation of Prediction by Linear Least Squares Methods*, 2nd edition, University of Minnesota Press, 1983.

Chapter 2

Introduction to Recursive Control and Prediction Theory

1. Introduction

In the text and problems of the preceding chapter, we described a class of discrete time opitmal control and filtering problems, described how to solve them by classical methods, and noted that the control and filtering problems had equivalent mathematical structures. Not surprisingly, the relationship between the two classes of problems again surfaces when recursive techniques are applied to these problems. By recursive techniques we mean the application of dynamic programming to the control problems, and of Kalman filtering to the filtering problems.

The purpose of this chapter is briefly to introduce the dynamic programming and the Kalman filtering algorithms, and to point out their formal equivalence. By pointing out their equivalence early on, we hope to double the reader's interest in the subsequent sections on controllability and reconstructibility. These concepts are of interest because it is in terms of them that conditions for the convergence and other important properties of the recursive algorithms are developed.

This chapter also contains a number of examples of control and filtering problems that have interested economists. We indicate how they fit into our framework.

The appendix contains statements of a few facts about linear least squares projections. Familiarity with Sargent [, Ch. 10] would also help the reader.

2. The Optimal Linear Regulator Control Problem

One problem that we shall study extensively is the *optimal linear regulator* problem. We consider a system with a $(n \times 1)$ *state* vector x_t and a $(k \times 1)$ control vector u_t. The system is assumed to evolve according to the law of motion

$$x_{t+1} = A_t x_t + B_t u_t \qquad t = t_0, t_0 + 1, \ldots, t_1 - 1,$$

where A_t is an $(n \times n)$ matrix and B_t is an $(n \times k)$ matrix. Both A_t and B_t are known sequences of matrices. We define the *return function* at time t, $r_t(x_t, u_t)$, as the quadratic form

$$r_t(x_t, u_t) = (x_t' u_t') \begin{bmatrix} R_t & W_t \\ W_t' & Q_t \end{bmatrix} \begin{pmatrix} x_t \\ u_t \end{pmatrix} \qquad t = t_0, \ldots, t_1 - 1$$

where R_t is $(n \times n)$, Q_t is $(k \times k)$ and W_t is $(n \times k)$. We shall initially assume that the matrices $\begin{pmatrix} R_t & W_t \\ W_t' & Q_t \end{pmatrix}$ are negative semi-definite, though subsequently we shall see that the problem can still be well-posed even if this assumption is weakened. We are also given an $(n \times n)$ negative semi-definite matrix P_t, which is a meteric for terminal values of the state x_{t_1}.

The *optimal linear regulator* problem is to maximize

$$\sum_{t=t_0}^{t_1} [x_t' u_t'] \begin{bmatrix} R_t & W_t \\ W_t' & Q_t \end{bmatrix} \begin{bmatrix} x_t \\ u_t \end{bmatrix} + t_{t_1}' P_{t_1} X_{t_1} \tag{2.1}$$

$$\text{subject to} \qquad x_{t+1} = A_t x_t + B_t u_t, \qquad x_{t_0} \text{ given.}$$

The maximization is carried out over the sequence of controls $(u_{t_0}, u_{t_0+1}, \ldots, u_{t_1-1})$. This is a recursive or serial problem, which it is appropriate to solve using the method of dynamic programming. In this case, the *value functions* are defined as the quadratic forms, $s = t_0, t_0 + 1, \ldots, t_1 - 1$,

$$x_s' P_s x_s = \max \left\{ \sum_{t=s}^{t_1} [x_t' u_t'] \begin{bmatrix} R_t & W_t \\ W_t' & Q_t \end{bmatrix} \begin{bmatrix} x_t \\ u_t \end{bmatrix} + x_{t_1}' P_{t_1} X_{t_1} \right\}$$

$$s.t. \; x_{t+1} = A_t x_t + B_t u_t,$$

x_s given $s = t_0, t_0 + 1, \ldots, t_1 - 1$. *Bellman's equation* becomes the following backward recursion in the quadratic forms $x_t P_t x_t$:

$$x_t' P_t x_t = \max_{u_t} \left\{ x_t' R_t x_t + u_t' Q_t u_t + 2x_t' W_t u_t + (Ax_t + Bu_t)' P_{t+1} (Ax_t + Bu_t) \right\},$$

$$t = t_1 - 1, t_1 - 2, \ldots, t_0$$

$$P_{t_1} \text{ given .}$$

Using the rules for differentiating quadratic forms (see appendix), the first-order necessary condition for the problem on the right side of (2.3) is found by differentiating with respect to the vector u_t:

$$\left\{ Q_t + B_t' P_{t+1} B_t \right\} u_t = -(B_t' P_{t+1} A_t + W_t') x_t.$$

Solving for u_t we obtain

$$u_t = -(Q_t + B_t' P_{t+1} B_t)^{-1} (B_t' P_{t+1} A_t + W_t') x_t. \tag{2.4}$$

The inverse $(Q_t + B_t'P_{t+1}B_t)^{-1}$ is assumed to exist. Otherwise, it could be interpreted as a generalized inverse, and most of our results would go through.

Equation (2.4) gives the optimal control in terms of a *feedback rule* upon the state vector x_t, of the form

$$u_t = -F_t x_t \tag{2.5}$$

where

$$F_t = (Q_t + B_t'P_{t+1}B_t)^{-1}(B_t'P_{t+1}A_t + W_t'). \tag{2.6}$$

Substituting (2.4) for u_t into (2.3) and rearranging gives the following recursion for P_t:

$$P_t = R_t + A_t'P_{t+1}A_t - (A_tP_{t+1}B_t + W_t)(Q_t + B_tP_{t+1}B_t)^{-1}(B_t'P_{t+1}A_t + W_t').$$

Equation (2.7) is a version of the *matrix Riccati difference equation.*

Equations (2.7) and (2.4) provide a recursive algorithm for computing the optimal controls in feedback form. Starting at time $(t_1 - 1)$, and given P_{t_1}, (2.4) is used to compute $u_{t_1-1} = -F_{t_1-1}x_{t_1-1}$. Then (2.7) is used to compute P_{t_1-1}. Then (2.4) is used to compute $u_{t_1-1} = F_{t_1-1}x_{t_1-1}$, and so on.

By substituting the optimal control $u_t = -F_t x_t$ into the state equation (2.1), we obtain the optimal *closed loop system* equations

$$x_{t+1} = (A_t - B_tF_t)x_t.$$

Eventually, we shall be concerned extensively with the properties of the optimal closed loop system, and how they are related to the properties of A, B, Q, and R.

3. Converting a Problem with Cross-Products in States and Controls to One With No Such Cross-Products

For our future work it is useful to introduce a problem that is equivalent with (2.1) - (2.2), and has a form in which no cross-products between states and controls appear in the objective function. This is useful because our theorems about the properties of the solutions (2.4) and (2.7) will be in terms of the special case in which $W = 0$. The equivalence between the problems (2.1) - (2.2) and the following problem implies that no generality is lost by restricting ourselves to the case in which $W = 0$.

The equivalent problem

$$\max_{\{u_t^*\}} \sum_{t=t_0}^{t_1} = \left\{x_t'(R_t - W_tQ_t^{-1}W_t')x_t + u_t^*Q_t'u_t^*\right\} + x_{t_1}'P_{t_1}x_{t_1} \tag{2.8}$$

subject to

$$x_{t+1} = (A_t - B_tQ_t^{-1}W_t')x_t + B_tu_t^* \tag{2.9}$$

and x_{t_0}, P_{t_0} are given. The new control variable u_t^* is related to the original control u_t by

$$u_t^* = Q_t^{-1}W_t'x_t + u_t. \tag{2.10}$$

We can state the problem (2.8) - (2.9) in a more compact notation as being to maximize

$$\sum_{t=t_0}^{t_1} \left\{x_t'\bar{R}_tx_t + u_t^{*\prime}Q_tu_t^*\right\} \tag{2.11}$$

subject to

$$x_{t+1} = \bar{A}_tx_t + B_tu_t^* \tag{2.12}$$

where

$$\bar{R}_t = R_t - W_tQ_t^{-1}W_t' \tag{2.13}$$

and

$$\bar{A}_t = A_t - B_tQ_t^{-1}W_t'. \tag{2.14}$$

With these specifications, the solution of the problem can be computed using the following versions of (2.4) and (2.7)

$$u_t^* = -\bar{F}x_t \equiv -(Q_t + B_t'P_{t+1}B_t)^{-1}B_tP_{t+1}\bar{A}_t \tag{2.15}$$

$$P_t = \bar{R}_t + \bar{A}_t'P_{t+1}\bar{A}_t - \bar{A}_t'P_{t+1}B_t(Q_t + B_t'P_{t+1}B_t)^{-1}B_t'P_{t+1}\bar{A}_t \tag{2.16}$$

We ask the reader to verify the following facts:

a. Problems (2.1) - (2.2) and (2.8) - (2.9) are equivalent.

b. The feedback laws $\bar{F}_t$ and F_t for u_t^* and u_t, respectively, are related by

$$F_t = \bar{F}_t + Q_t^{-1} W_t'.$$

c. The "closed loop" transition matrices are related by

$$A_t - B_t F_t = \bar{A}_t - B_t \bar{F}_t.$$

4. An Example

We now give an example of a problem for which the preceding transformation is useful. A consumer wants to maximize

$$\sum_{t=t_0}^{\infty} \beta^t \left\{ u_1 c_t - \frac{u_2}{2} c_t^2 \right\} \quad 0 < \beta 1 \tag{2.17}$$

subject to the intertemporal budget constraint

$$k_{t+1} = (1+r)\,[k_t + y_t - c_t], \tag{2.18}$$

the law of motion for labor income

$$y_{t+1} = \lambda_0 + \lambda_1 y_t, \tag{2.19}$$

and a given level of initial assets, k_{t_0}. Here β is a discount factor, c_t is consumption, k_t is "nonhuman" assets at the beginning of time t, $r > -1$ is the interest rate on nonhuman assets, and y_t is income from labor at time t.

We define the transformed variables

$$\bar{k}_t = \beta^{t/2} k_t$$
$$\bar{y} = \beta^{t/2} y_t$$
$$\bar{c}_t = \beta^{t/2} c_t.$$

In terms of these transformed variables, the problem can be re-written as follows: maximize

$$\sum_{t=t_0}^{\infty} \left\{ u_1 \beta^{t/2} \cdot \bar{c}_t - \frac{u_2}{2} \bar{c}_t^2 \right\} \tag{2.20}$$

subject to

$$(2.21)\qquad \begin{aligned} \bar{k}_{t+1} &= (1+\lambda)\beta^{1/2}[\bar{k}_t + \bar{y}_t - \bar{c}_t] \quad \text{and} \\ \bar{y}_{t+1} &+ \lambda_0\beta^{\frac{t+1}{2}} + \lambda_1\beta^{1/2}\bar{y}_t \end{aligned}$$

and k_{t_0} given. We write this problem in the state-space form:

$$\max_{\{\bar{u}_t\}} \sum_{t=t_0}^{\infty} \left\{\bar{x}_t' R\bar{x}_t + 2\bar{x}_t' W\bar{u}_t + \bar{u}_t' Q\bar{u}_t\right\}$$
$$s.t.\ \bar{x}_{t+1} = A\bar{x}_t + B\bar{u}_t.$$

We take

$$\bar{x}_t = \begin{bmatrix} \bar{x}_t \\ \bar{y}_t \\ \beta^{t/2} \end{bmatrix},\ \bar{u}_t = \bar{c}_t,$$

$$R = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},\ W' = [0\,0\,\frac{u_1}{2}],$$

$$Q = -\frac{u_2}{2},\ A = \begin{bmatrix} (1+r) & (1+r) & (1+r) \\ 0 & \lambda_1 & \lambda_0 \\ 0 & 0 & 1 \end{bmatrix}\beta^{1/2},\ B = \begin{bmatrix} -(1+4) \\ 0 \\ 0 \end{bmatrix}\beta^{1/2}.$$

To obtain the equivalent transformed problem in which there are no cross-product terms between states and controls in the return function, we take

$$\bar{A} = A - BQ^{-1}W' = \begin{bmatrix} (1+r) & (1+r) & -\frac{u_1 d(1+r)}{u_2} \\ 0 & \lambda_1 & \lambda_0 \\ 0 & 0 & 1 \end{bmatrix}\beta^{1/2}$$

$$(2.22)\qquad \bar{R} = R - WQ^{-1}W' = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{2u_1^2}{u_2} \end{bmatrix}$$

$$U^* + t = \bar{u}_t + Q^{-1}W'\bar{x}_y$$

$$c_t^* = \bar{c}_t - \frac{u_1}{u_2}\beta^{t/2}$$

Thus, our original problem can be expressed as:

$$(2.23)\qquad \max_{\{u_t^*\}} \sum_{t=t_0}^{\infty} \left\{\bar{x}_t'\bar{R}\bar{x}_t + u_t^{*\prime} Q u_t^*\right\}$$
$$s.t.\ \bar{x}_{t+1} = \bar{A}\bar{x}_t + Bu_t^*.$$

For future reference, it will be useful to write problem (2.23) - (2.24) in the partitioned form:

$$\max_{\{u_t^*\}} \sum_{t=t_0}^{\infty} \left\{[\bar{x}_{1t}\bar{x}_{2t}] \begin{bmatrix} \bar{R}_{11} & \bar{R}_{12} \\ \bar{R}_{12}' & \bar{R}_{22} \end{bmatrix} \begin{bmatrix} \bar{x}_{1t} \\ \bar{x}_{2t} \end{bmatrix}\right\} + u_t^* Q u_t^*$$

subject to

$$\begin{bmatrix} \bar{x}_{1(t+1)} \\ \bar{x}_{2(t+1)} \end{bmatrix} = \begin{bmatrix} \bar{A}_{11} & \bar{A}_{12} \\ 0 & \bar{A}_{22} \end{bmatrix} \begin{bmatrix} \bar{x}_{1(t)} \\ \bar{x}_{2(t)} \end{bmatrix} + \begin{bmatrix} B_1 \\ 0 \end{bmatrix} u_t^*.$$

Here the partitional vectors and matrices are given by

$$\bar{x}_{1(t)} = \bar{k}_{(t)}$$

$$\bar{x}_{2(t)} = \begin{bmatrix} \bar{y}_{(t)} \\ \beta^{t/2} \end{bmatrix}$$

$$\bar{A} = \begin{bmatrix} (1+r) & (1+r) & \frac{u_1}{u_2} & (1+r) \\ 0 & & \lambda_1 & \lambda_2 \\ 0 & & 0 & 1 \end{bmatrix} \beta^{1/2}$$

$$\bar{R} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{u_1^2}{u_2} \end{bmatrix} \beta^{1/2}$$

$$B = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = \begin{bmatrix} -(1+r) \\ 0 \\ 0 \end{bmatrix}.$$

Notice that the pattern of 0's in $\bar{A}$ and B, and in particular that $\bar{A}_{12} = 0$ and $B_2 = 0$. Later on we shall be concerned extensively with properties of linear spaces generated by the certain functions of the pair of matrices $(\bar{A}_{11}, R_{11})$ and the pair $(\bar{A}_{11}, B_1)$.

5. The Kalman Filter

We consider the linear system

$$x_{t+1} = A_t x_t + B_t u_t + G_t w_{1t+1} \tag{2.24}$$

$$y_t = C_t x_t + E_t u_t + w_{2t} \tag{2.25}$$

where $[w'_{1t+1} \cdot w'_{2t}]$ is a vector white noise with contemporaneous covariances matrix

$$E \begin{bmatrix} w_{1t+1} \\ w_{2t} \end{bmatrix} \begin{bmatrix} w_{1t+1} \\ w_{2t} \end{bmatrix}' = \begin{bmatrix} V_{1t} & V_{3t} \\ V_{3t} & V_{2t} \end{bmatrix} \geq 0.$$

The w_{1t+1}, w_{2t}, vector for $t \geq t_0$ is assumed orthogonal to the initial condition x_{t_0}, which represents the analyst's initial ideas about the state. Here, A_t is $(n \times n)$, B_t is $(n \times k)$, G_t is $(n \times n)$, C_t is $(\ell \times n)$, E_t is $(\ell \times n)$, w_{1t+1} is $(n \times 1)$, w_{2t+1} is $(\ell \times 1)$; x_t is an $(n \times 1)$ vector of *state* variables, u_t is a $(k \times 1)$ vector of *controls*, and y_t is an $(\ell \times 1)$ vector of *output* or observed variables. The matrices A_t, B_t, G_t, C_t, and E_t are known, though possibly time varying. The noise vector w_{1t+1} is the state-disturbance, while w_{2t} is the measurement error.

The analyst does not directly observe the x_t process. So from his point of view, x_t is a "hidden state vector". The system is assumed to start up at time t_0, at which time the state vector x_{t_0} is regarded as a random variable with mean $Ex_{t_0} = \hat{x}_{t_0}$, and given covariance matrix $\sum_{t_0} = \sum_0$. The pair $(\hat{x}_{t_0}, \sum_0)$ can be regarded as the mean and covariance of the analyst's Bayesian prior distribution on x_{t_0}.

It is assumed that for $s \geq 0$, the vector of random variables $\left[\begin{smallmatrix} w_{1t_0+s+1} \\ w_{2t_0+s} \end{smallmatrix}\right]$ is orthogonal to the random variable x_{t_0} and to the random variables $\left[\begin{smallmatrix} w_{1t_0+r+1} \\ w_{2t_0+r} \end{smallmatrix}\right]$ for $r \neq s$. It is also assumed that $E\left[\begin{smallmatrix} w_{1t_0+s+1} \\ w_{2t+s} \end{smallmatrix}\right] = 0$ for $s \geq 0$. Thus, $\left[\begin{smallmatrix} w_1(t+1) \\ w_2(t) \end{smallmatrix}\right]$ is a serially uncorrelated or white noise process. Further, from (2.24), (2.25) and the orthogonality properties posited for $\left[\begin{smallmatrix} w_{1t+1} \\ w_{2t} \end{smallmatrix}\right]$ and x_{t_0}, it follows that $\left[\begin{smallmatrix} w_{1t+1} \\ w_{2t} \end{smallmatrix}\right]$ is orthogonal to $\{x_s, y_{s-1}\}$ for $s \leq t$. This follows because y_t and x_{t+1} are in the space spanned by current and lagged u_t, w_{1t+1}, w_{2t}, and x_{t_0}.

The analyst is assumed to observe at time t $\{y(s), u(s) \,:\, s = t_0, t_0+1, \ldots t\}$, for $t = t_0, t_0+1, \ldots t_1$. The object is then to compute the linear least squares projection of the state x_{t+1} on this information, which we denote $\widehat{E}_t x_{t+1}$. We write this projection as

$$\widehat{E}_t x_{t+1} \equiv \widehat{E}[x_{t+1} \mid y_t, y_{t+1}, \ldots, y_{t_0}, \hat{x}_{t_0}] \tag{2.26}$$

where $\hat{x}_{t_0}$ is the initial estimate of the state. It is convenient to let Y_t denote the information on y_t collected through time t:

$$Y_t = \{y_t, y_{t-1}, \ldots, y_{t_0}\}.$$

The linear least squares projection of y_{t+1} on Y_t, and $\hat{x}_{t_0}$ is from (2.25) and (2.26) given by

$$\begin{aligned} \widehat{E}_t y_{t+1} &\equiv \widehat{E}(y_{t+1}) \mid Y_t, \hat{x}_{t+1} \\ &= C_t \widehat{E}_t x_{t+1} + E_t\, u_{t+1} \end{aligned} \tag{2.27}$$

since w_{2t+1} is orthogonal to $\{w_{1s+1},\, w_{2s}\}$, $s \leq t$, x_{t_0} and is therefore orthogonal to $\{Y_t,\, x_{t_0}\}$.

In the interests of conveniently constructing the projections $\widehat{E}_t x_{t+1}$ and $\widehat{E}_t y_{t+1}$, we now apply a Gram-Schmidt orthogonalization procedure to the set of random variables $\{x_{t_0}, y_{t_0}, y_{t_0+1}, \ldots y_{t_1}\}$. An orthogonal basis for this set of random variables is formed by the set $\{\hat{x}_{t_0}, \tilde{y}_{t_0}\tilde{y}_{t_0+1}, \dots, \tilde{y}_{t_1}\}$ where

$$\tilde{y}_t = y_t - \widehat{E}[y_t \mid \tilde{y}_{t-1}, \tilde{y}_{t-2}, \dots \tilde{y}_{t_0}, \hat{x}_{t_0}]. \tag{2.28}$$

For convenience, let us write $\tilde{Y}_t = \{\tilde{y}_{t_0}, \tilde{y}_{t_0+1}, \dots, \tilde{y}_t\}$. We note that the linear spaces spanned by $(\hat{x}_{t_0}, Y_t)$ equals the linear space spanned by $(\hat{x}_{t_0}, \tilde{Y}_t)$. This follows because: (a) $\tilde{y}_t$ is formed as indicated above as a linear function of Y_t, and $\hat{x}_{t_0}$, and (b) y_t can be recovered from $\tilde{Y}_t$ by noting that $y_t = E[y_t \mid x_{t_0}, \tilde{Y}_{t-1}] + \tilde{y}_t$. It follows that $\widehat{E}[y_t \mid x_{t_0}, Y_{t-1}] = \widehat{E}[y_t \mid x_{t_0}, \tilde{Y}_{t-1}] = E_{t-1}y_t$. In (2.28), we use (2.25) to write

$$\widehat{E}[y_t \mid \hat{x}_{t_0}] = C_{t_0}\hat{x}_{t_0} + E_{t_0}u_{t_0}.$$

To summarize developments up to this point, we have defined the *innovation process*

$$\begin{aligned}
\tilde{y}_t &= y_t - \widehat{E}[y_t \mid \hat{x}_{t_0}, Y_{t-1}] \\
&= y_t - \widehat{E}[y_t \mid \hat{x}_{t_0}, \tilde{Y}_{t-1}], \ t \geq t_0 + 1 \\
\tilde{y}_{t_0} &= y_{t_0} - \hat{E}[y_{t_0} \mid \hat{x}_{t_0}].
\end{aligned}$$

The innovations process is *serially uncorrelated* ($\tilde{y}_t$ is orthogonal to $\tilde{y}_s$ for $t \neq s$) and spans the same linear space as the original Y process.

We now use the innovations process to get a recursive procedure for evaluating $\widehat{E}_t x_{t+1}$. Using theorem $A3$ about projections on orthogonal bases gives

$$\begin{aligned}
&\widehat{E}[x_{t+1} \mid \hat{x}_{t_0}, \tilde{y}_{t_0}, \tilde{y}_{t_0+1}, \dots, \tilde{y}_t] \\
&\quad = \widehat{E}x_{t+1} \mid \tilde{y}_t] + \widehat{E}[x_{t+1} \mid x_{t_0}, \tilde{y}_{t_0}, \tilde{y}_{t_0+1}, \dots, \tilde{y}_{t-1}] - Ex_{t+1}
\end{aligned} \tag{2.29}$$

We have to evaluate the first two terms on the right-hand side of (2.29).

From theorem $A1$, we have that

$$\widehat{E}[x_{t+1} \mid \tilde{y}_t] = Ex_{t+1} + \text{cov } (x_{t+1}, \tilde{y}_t) \ [\text{cov } (\tilde{y}_t, \tilde{y}_t)]^{-1}\tilde{y}_t. \tag{2.30}$$

To evaluate the covariances that appear in (2.30), we shall use the covariance matrix of one-step ahead errors, $\tilde{x}_{(t)} = x_{(t)} - \widehat{E}_{t-1}x_t$, in estimating x_t. We define this covariance matrix as $\sum_t = E\tilde{x}_t\tilde{x}_t'$. It follows from (2.24) and (2.25) that

(2.31)

$$\begin{aligned}
\operatorname{cov}(x_{t+1}, \tilde{y}_t) &= \operatorname{cov}(A_t x_t + B_t u_t - G_t w_{1t+1}, y_t - \widehat{E}_{t-1} y_t) \\
&= \operatorname{cov}(A_t x_t + B_t u_t + G_t w_{1t+1}, C_t x_t + w_{2t} - c_t \widehat{E}_{t-1} x_t) \\
&= \operatorname{cov}(A_t x_t + B_t u_t + G_t w_{1t+1}, C_t \tilde{x}_t + w_{2t}) \\
&= E[(A_t x_t + B_t u_t + G_t w_{1t+1} - E(A x_t + B u_t + G_t w_{1t+1})(C_t \tilde{x}_t + w_{2t} - E(C_t \tilde{x}_t + w_{2t}))] \\
&= E[(A_t x_t + G_t w_{1t+1} - A_t E x_t)(\tilde{x}_t' C_t' + w_{2t}')] \\
&= E[A_t x_t \tilde{x}_t' C_t'] + G_t E[w_{1t+1} \tilde{x}_t' C_t'] - A_t E x_t E \tilde{x}_t' C_t' + A_t E[x_t w_{2t}'] \\
&\qquad + G_t E[w_{1t+1} w_{2t}'] - A_t E x_t E w_{2t}' \\
&= E[A_t x_t \tilde{x}_t' C_t'] + G_t E[w_{1t+1} w_{2t}'] \\
&= E[A_t(\tilde{x}_t - \widehat{E}_{t-1} x_t)\tilde{x}_t' C_t'] + G_t E[w_{1t+1} w_{2t}'] \\
&= A_t E \tilde{x}_t \tilde{x}_t' C_t' + G_t E(w_{1t+1} w_{2t}' = A_t \Sigma_t C_t' + G_t V_{3t}
\end{aligned}$$

The second equality was the fact that $\widehat{E}_{t-1} w_{2t} = 0$ since w_{2t} is orthogonal to $\{x_s,\ y_{s-1}\}$, $s \leq t$. To get the fifth equality, we use the fact that $E\tilde{x}_t = E[x_t - \widehat{E}_{t-1} x_t] = 0$ by the unbiased property of linear projections. We also use the facts that u_t is known and w_{1t+1} and w_{2t} have zero means. The seventh equality follows from the orthogonality of w_{1t+1} and w_{2t} to variables dated t and earlier and the means of w_{2t}' and $\tilde{x}_t'$ being zero. Finally, the eighth equation relies on the fact that $\tilde{x}_t$ is orthogonal to the subspace generated by $y_{t-1}, y_{t-2}, \ldots$ and $\widehat{E}_{t-1} x_t$ is a function of these vectors.

Next we evaluate

$$\begin{aligned}
\operatorname{cov}(\tilde{y}_t, \tilde{y}_t) &= E(C_t \tilde{x}_t + w_{2t})(c_t \tilde{x}_t + w_{2t})' \\
&= C_t \Sigma_t C_t' + V_{3t}
\end{aligned}$$

and since $E\tilde{y}_t = 0$ and $E\tilde{x}_t w_{2t}' = 0$. Therefore, (2.31) becomes

$$E[x_{t+1} \mid \tilde{y}_t] = E[x_{t+1}] + (A_t \Sigma_t C_t' + G_t V_{3t})(C_t \Sigma_t C_t' + V_{2t})^{-1} \tilde{y}_t. \tag{2.32}$$

Using equation (2.24), we evaluate the second term on the right side of (2.29),

$$E[x_{t+1} \mid \tilde{Y}_{t-1}, \hat{x}_{t_0}] = A_t E[x_t \mid \tilde{Y}_{t-1}, \hat{x}_{t_0}] + B_t u_t$$

or

$$\widehat{E}_{t-1} x_{t+1} = \widehat{E}_{t-1} x_t + B_t u_t. \tag{2.33}$$

Using (2.32) and (2.33) in (2.29) gives

$$\widehat{E}_t x_{t+1} = A_t \widehat{E}_{t-1} x_t + B_t u_t + K_t(y_t - E_{t-1} y_t) \tag{2.34}$$

where

$$K_t = \left(A_t \Sigma_t C_t' + G_t V_{3t}\right)\left(C_t \Sigma_t C_t' + V_{2t}\right)^{-1}. \tag{2.35}$$

Using $\widehat{E}_{t-1} y_t = C_t \widehat{E}_{t-1} x_t + E_t u_t$, equation (2.34) can also be written

$$\widehat{E}_t x_{t+1} = [A_t - K_t C_t]\widehat{E}_{t-1} x_t + [B_t - K_t E_t] u_t + K_t y_t.$$

We now aim to derive a recursive formula for the covariance matrix Σ_t. From equation (2.25) we have that $\widehat{E}_{t-1} y_t = C_t \widehat{E}_{t-1} x_t + E_t u_t$. Subtracting this from y_t in (2.25) gives

$$y_t - E_{t-1} y_t = C_t[x_t - E_{t-1} x_t] + w_{2t}.$$

Substituting this expression in (2.34) and subtracting the result from (2.24) gives

$$\begin{aligned}(x_{t+1} - \widehat{E}_t x_{t+1}) = (A_t - K_t C_t)(x_t - \widehat{E}_{t-1} x_t)\\ + G_t w_{1t+1} - K_t w_{2t}\end{aligned}$$

or

$$\tilde{x}_{t+1} = (A_t - K_t C_t)\tilde{x}_t + G_t w_{1t+1} - K_t w_{2t}. \tag{2.36}$$

From (2.36) and our specification of the covariance matrix

$$E\begin{bmatrix} w_{1t+1} \\ w_{2t} \end{bmatrix}\begin{bmatrix} w_{1t+1} \\ w_{2t} \end{bmatrix}' = \begin{bmatrix} V_{1t} & V_{3t} \\ V_{3t}' & V_{2t} \end{bmatrix}$$

we have

$$\begin{aligned} E\tilde{x}_{t_1}\tilde{x}_{t_1}' = {} & \left[A_t - K_t C_t\right] E\tilde{x}_t \tilde{x}_t' \left[A_t - K_t C_t\right]' \\ & + G_t V_{1t} G_t' + K_t V_{2t} K_t \\ & - G_t V_{3t} K_t' - K_t V_{3t}' G_t' \end{aligned}$$

We have defined the covariance matrix of $\tilde{x}_t$ as $\Sigma_t = E\tilde{x}_t\tilde{x}_t' = E(x_t - E_{t-1}x_t)(x_t - E_{t-1}x_t)'$. So we can express the above equation as

$$\begin{aligned} \Sigma_{t+1} = {} & \left[A_t - K_t C_t\right]\Sigma_t\left[A_t - K_t C_t\right]' \\ & + G_t V_{1t} G_t' + K_t V_{2t} K_t' - G_t V_{3t} K_t' \\ & - K_t V_t' G_t. \end{aligned} \tag{2.37}$$

Equation (2.37) can be rearranged to the equivalent form

$$\begin{aligned}\Sigma_{t+1} &= A_t \Sigma_t A_t' + G_t V_{1t} G_t' \\ &\quad - \left[A_t \Sigma_t C_t' + G_t V_{3t}\right]\left(C_t \Sigma_t C_t' + V_{2t}\right)^{-1} \\ &\quad \times \left(A_t \Sigma_t' + G_t V_{3t}\right)'\end{aligned}$$

We repeat (2.35) here for your convenience

$$K_t = \left(A_t \Sigma_t C_t' + G_t V_{3t}\right)\left(C_t \Sigma_t C_t' + V_{2t}\right)^{-1} \tag{2.35}$$

Starting from the given initial condition for $\Sigma_{t_0} = E(x_{t_0} - Ex_{t_0})(x_{t_0} - Ex_{t_0})'$, equations (2.37) and (2.35) give a recursive procedure for generating the "Kalman gain" K_t, which is the crucial unknown ingredient of the recursive algorithm (2.34) for generating $\widehat{E}_t x_{t+1}$.

The Kalman filter is used as follows. Starting from time t_0 with $\Sigma_{t_0} = \Sigma_0$ and $\hat{x}_{t_0} = x_0$ given, (2.35) is used to form K_{t_0}, and (2.34) is used to obtain $\widehat{E}_{t_0} x_{t_0+1}$ with $E_{t_0-1} x_{t_0} = \hat{x}_0$. Then (2.37) or (2.38) is used to form Σ_{t_0+1}, (2.35) is used to form K_{t_0+1}, (2.34) is used to obtain $\widehat{E}_{t_0+1} x_{t_0+2}$, and so on.

The evolution of the state estimate obeys

$$\hat{x}_{t+1} = (A_t - K_t C_t)\hat{x}_{t+1} + K_t y_t \tag{2.39}$$

where

$$y_t = C_t x_t + w_{2t}. \tag{2.40}$$

We can represent y_t as

$$y_t = C_t \hat{x}_t + a_t \tag{2.41}$$

where

$$a_t = w_{2t} + C_t(x_t - \hat{x}_t) \tag{2.42}$$

Now from (2.40) it follows that

$$\hat{y}_t \equiv \widehat{E}_{t-1} y_t = C_t \hat{x}_t \tag{2.43}$$

Therefore from (2.41) and (2.43) we have

$$y_t - \hat{y}_t = C_t(x_t - \hat{x}_t) + w_{2t}$$

or

$$y_t - \hat{y}_t = a_t.$$

We have that $Ea_t a_t' = C_t \Sigma_t C_t' + V_{2t}$. The random process a_t is the "innovation" in y_t, i.e. the part of y_t that cannot be predicted linearly from past y's.

Using (2.41) the system (2.39)–(2.40) can be presented as

$$\begin{aligned} \hat{x}_{t+1} &= A_t \hat{x}_t + K_t a_t \\ y_t &= C_t \hat{x}_t + a_t \end{aligned} \tag{2.44}$$

System (2.44) is called an "innovations representation."

Another representation of the system which is useful is obtained by combining (2.39) with (2.41) to get

$$\begin{aligned} \hat{x}_{t+1} &= (A_t - K_t C_t)\hat{x}_t + K_t y_t \\ a_t &= y_t - C_t \hat{x}_t \end{aligned} \tag{2.45}$$

This is called a "whitening filter." Starting from a given $\hat{x}_0$, this system accepts as an "input" a history of y_t and gives as an output the sequence of innovations a_t, which by construction are serially uncorrelated.

We shall often study situations in which the system is time invariant, i.e. $A_t = A, C_t = C, V_{jt} = V_j$ for all t. We shall later describe regulatory conditions on A, C, V_1, V_2, V_3 which imply that (i) $K_t \to K$ as $t \to \infty$ and $\Sigma_t \to \Sigma$ as $t \to \infty$; and (ii) $\mid \lambda_i(A - KC) \mid < 1$ for all i, whose λ_i is the ith eigesvalue of $(A - KC)$. When these conditions are met, the limiting representation for (2.44) is time invariant and is an (infinite dimensional) innovations representation. Using the lag operator L where by $L\hat{x}_t = \hat{x}_t$, imposing time invariance in (2.44) and rearranging gives the representation

$$y_t = [I + C(L^{-1}I - A)^{-1}K]a_t \tag{2.46}$$

which expresses y_t as a function of $[a_t, a_{t-1}, \ldots]$. In order that $[y_t, y_{t-1}, \ldots]$ span the same linear space as $[a_t, a_{t-1}, \ldots]$, it is necessary that the following condition be met:

$$det\,[I + C(z - A)^{-1}K] = 0 \;\Rightarrow\; \mid z \mid < 1.$$

Now by a theorem in linear algebra we have that

$$det[I + C(zI - A)^{-1}K] = \frac{det(zI - (A - KC))}{det(zI - A)}.$$

The formula shows that the zeros of $det[I + C(zI - A)^{-1}K]$ equal that zeros of $det(zI - (A - KC))$, which are the eigenvalues of $A - KC$. Thus, if the eigenvalues of $(A - KC)$ are all less than unity in modulus, then the spaces $[a_t, a_t, \ldots]$ and $[y_t, y_{t-1}, \ldots]$ in representation (2.46) are equal.

6. Duality

For purposes of highlighting their relationship, we now repeat the Kalman filtering formulas for K_t and Σ_t and the optimal linear regulator formulas for F_t and P_t

$$K_t = \left(A_t\Sigma_tC_t' + G_tV_{3t}\right)\left(C_t\Sigma_tC_t' + V_{2t}\right)^{-1}. \tag{2.35}$$

$$\begin{aligned}\Sigma_{t+1} &= A_t\Sigma_tA_t' + G_tV_{1t}G_t' \\ &\quad - \left(A_t\Sigma_tC_t' + G_tV_{3t}\right)\left(C_t\Sigma_tC_t' + V_{2t}\right)^{-1} \\ &\quad \times \left(A_t\Sigma_tC_t' + G_tV_{3t}\right)'\end{aligned}$$

$$F_t = (Q_t + B_t'P_{t+1}B_t)^{-1}(B_t'P_{t+1}A_t + W_t'). \tag{2.6}$$

$$\begin{aligned}P_t &= R_t + A_t'P_{t+1}A_t \\ &\quad - (A_tP_{t_1}B_t + W_t')(Q_t + B_t'P_{t+1}B_t)^{-1} \\ &\quad \times \left(B_t'P_{t+1}B_t + W_t\right)\end{aligned} \tag{2.7}$$

for $t = t_0, t_0 + 1, \ldots, t_1$. The equations in (2.35) are solved forwards from t_0 with Σ_{t_0} given while those in (2.6) and (2.7) are solved backwards from $t_1 - 1$ with P_{t_1} given.

Table 1

Object in Optimal Linear Regulator Problem	Object in Kalman Filter
$A_{t_0+s}, s = 0, \ldots, t_1 - t_0 - 1$	$A'_{t_1-1-s}, s = 0, \ldots, t_1 - t_0 - 1$
B_{t_0+s}	C'_{t_1-1-s}
R_{t_0+s}	$-G_{t_1-1-s} V_{1t_1-1-s} G'_{t_1-1-s}$
Q_{t_0+s}	$-V_{2t_1-1-s}$
W_{t_0+s}	$-G_{t_1-1-1} V_{3t_1-1-s}$
P_{t_0+s}	$-\Sigma_{t_1-s}$
F_{t_0+s}	K'_{t_1-1-s}
P_{t_1}	$-\Sigma_{t_0}$
$A_{t_0+s} - B_{t_0+s} F_{t_0+s}$	$A'_{t_1-1-s} - C'_{t_1-1-s} K'_{t_1-1-s}$

The equations for K_t and F_t are intimately related, as are the equations for P_t and Σ_t. In fact, upon properly re-interpreting the various matrices in (2.35), (2.6) and (2.7), the equations for the Kalman filter and the optimal linear regulator can be seen to be identical. Thus, where A appears in the Kalman filter, A' appears in the corresponding regulator equation, where C appears in the Kalman filter, B' appears in the corresponding regulator equation, and so on. The correspondences are listed in detail in Table 1. By taking account of these correspondences, a single set of computer programs can be used to solve either an optimal linear regulator problem or a Kalman filtering problem.

The concept of *duality* helps to clarify the relationship between the optimal regulator and the Kalman filtering problem.

Definition 2.1: Consider the time varying linear system.

$$
\begin{aligned}
x_{t+1} &= A_t x_t + B_t u_t \\
y_t &= C_t x_t, \; t = t_0, \ldots, t_1 - 1
\end{aligned} \tag{2.47}
$$

The *dual* of system (2.47) (sometimes called the "dual with respect to $t_1 - 1$") is the system:

$$
\begin{aligned}
x^*_{t+1} &= A'_{t_1-1-t} x^*_t + C'_{t_1-1-t} u^*_t \\
y^*_t &= B'_{t_1-1-t} x^*_t
\end{aligned}
$$

with $t = t_0, t_0 + 1, \ldots, t_1 - 1$.

With this definition, the correspondence exhibited in Table 1 can be summarized succinctly in the following proposition:

Proposition 2.1: Let the solution of the optimal linear regulator problem defined by the given matrices $\{A_t, B_t, R_t, Q_t, W_t; t = t_0, \ldots, t_1 - 1; P_{t_1}\}$ be given by $\{P_t, F_t, \; t = t_0, \ldots, t_1 - 1\}$. Then the solution of the Kalman filtering problem definded by the matrices $\{A'_{t_1-1-t}, C'_{t_1-1-t}, \; -G_{t_1-1-t}V_{1t_1-1-t}, G_{t_1-1-t} - V_{2t_1-1-t}, -G_{t_1-1-t}V_{3t_1-1-t}; \; t = t_0, \ldots, t_1 - 1; \; \Sigma_{t_0}\}$ is given by $\{K'_{t_1-t-1} = F_t, -\Sigma_{t_1-t} = P_t; \; t = t_0, \, t_0 + 1, \ldots, t_1 - 1\}$.

This proposition describes the sense in which the Kalman filtering problem and the optimal linear regulator problems are "dual" to one another. As we also saw in our discussion of classical control and filtering methods, the very same equations arise in solving the filtering problem as arise in solving the control problem. This fact implies that most everything that we learn about the control problem applies to the filtering problem, and *vice versa*.

As an example of the use of duality, recall the transformations (2.13) and (2.14) that we used to convert the optimal linear regulator problem with cross-products between *states* and *controls* into an equivalent problem with no such cross-products. The preceding discussion of duality and Table 1 suggest that the same transformation will convert the original dual filtering problem which has nonzero covariance matrix V_3 between *state noise* and *measurement noise* into an equivalent problem with covariances zero. This hunch is correct. The transformations, which can be obtained by duality directly from (2.13)-(2.14), are for $t = t_0, \ldots, t_1 - 1$:

$$\bar{A}'_{t_1-1-t} = A'_{t_1-1-t} - C'_{t_1-1-t}V^{-1}_{2t_1-1-t}V'_{3t_1-1-t}G'_{t_1-1-t}$$

$$-\bar{V}_{1t_1-1-t} = -V_{1t_1-1-t} + V_{3t_1-1-t}V^{-1}_{2t_1-1-t}V'_{3t_1-1-t}$$

The Kalman filtering problem defined by the matrices $\{\bar{A}_t, C_t, -G_t\bar{V}_{1t} - V_{2t}, 0; \; t = t_0, \ldots, t_1 - 1; \; \Sigma_0\}$ is equivalent to the original problem in the sense that

$$A_t - K_tC_t = \bar{A}_t - \bar{K}_tC_t$$

where $\bar{K}_t$ is the solution of the transformed problem. We also have, by the results for the regulator problem and duality, that

$$\bar{K}_t = K_t - G_tV_{3t}V_{2t}^{-1}.$$

7. Examples of Kalman Filtering

This section contains several examples which have been widely used by economists and that fit into the Kalman filtering setting. After the reader has worked through our examaples, no doubt many other examples will recur to her or him.

a. Vector autoregression: We consider an $(n \times 1)$ stochastic process y_t that obeys the linear stochastic difference equation

$$y_t = A_1 y_{t-1} + \ldots + A_m y_{t-m} + \varepsilon_t$$

where ε_t is an $(h \times 1)$ vector white noise, with mean zero and $E\varepsilon_t \varepsilon_t' = V_{1t}$, $E\varepsilon_t y_s' = 0,\ t > s$. We define the state vector x_t and shock vector w_t as

$$x_t = \begin{bmatrix} y_{t-1} \\ y_{t-2} \\ \vdots \\ y_{t-m} \end{bmatrix}, \quad \begin{bmatrix} w_{1t+1} \\ w_{2t} \end{bmatrix} = \begin{pmatrix} \varepsilon_t \\ \varepsilon_t \end{pmatrix}.$$

The law of motion of the system then becomes

$$\begin{bmatrix} y_t \\ y_{t-1} \\ y_{t-2} \\ \vdots \\ y_{t-m+1} \end{bmatrix} = \begin{bmatrix} A_1 & A_2 & \ldots & A_m \\ I & 0 & \ldots & 0 \\ 0 & I & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & I & 0 \end{bmatrix} \begin{pmatrix} y_{t-1} \\ y_{t-2} \\ y_{t-3} \\ \vdots \\ y_{t-m} \end{pmatrix} = \begin{bmatrix} I \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \varepsilon_t.$$

The measurement equation is

$$y_t = [A_1\ A_2 \ldots A_m]\, x_t + \varepsilon_t.$$

For the filtering equations, we have

$$A_t = \begin{bmatrix} A_1 & A_2 & \ldots & A_m \\ I & 0 & \ldots & 0 \\ 0 & I & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & I & 0 \end{bmatrix}, \quad G_t = G = \begin{bmatrix} I \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$C_t = [A_1, \ldots, A_n]$$

$$V_{1t} = V_{2t} = V_{3t}.$$

Starting from $\Sigma_{t_0} = 0$, which means that the system is imagined to start up with m lagged values of y having been observed, (5.35) imples

$$K_{t_0} = G,$$

while (.16) implies that $\Sigma_{t_0+1} = 0$. It follows recursively that $K_t = G$ for all $t \geq t_0$ and that $\Sigma_t = 0$ for all $t \geq t_0$. Computing $(A - KC)$, we find that

$$\widehat{E}_t x_{t+1} = \begin{bmatrix} 0 & 0 & \dots & 0 \\ I & 0 & \dots & 0 \\ 0 & I & \dots & 0 \\ \vdots & & & \\ 0 & \dots & I & 0 \end{bmatrix} \widehat{E}_{t-1} x_t + \begin{bmatrix} I \\ 0 \\ \vdots \\ 0 \end{bmatrix} y_t,$$

which is equivalent with

$$\widehat{E}_t x_{t+1} = \begin{bmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-m} \end{bmatrix}.$$

The equation $\widehat{E}_t y_{t+1} = C \widehat{E}_t x_{t+1}$ becomes

$$\widehat{E}_t y_{t+1} = A_1 y_t + A_2 y_{t-1} + \dots A_m y_{t-m+1}.$$

Evidently, the preceding equation for forecasting a vector autoregressive process can be obtained in a much less roundabout manner, with no need to use the Kalman filter.

b. Univariate moving average: We consider the model

$$y_t = w_t + c_1 w_{t-1} + \dots + c_n w_{t-n}$$

where w_t is a univariate white noise with mean zero and variance V_{1t}. We write the model in the state-space form

$$x_{t+1} = \begin{bmatrix} w_t \\ w_{t-1} \\ \vdots \\ w_{t-n+1} \end{bmatrix} = \begin{bmatrix} 0 & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 1 & 0 \end{bmatrix} \begin{bmatrix} w_{t-1} \\ w_{t-2} \\ \vdots \\ w_{t-n} \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} w_t$$

$$y_t = [c_1\ c_2 \dots c_n] x_t + w_t.$$

We assume that $\Sigma_{t_0} = 0$, so that the initial state is known. In this setup, we have A, G and C as indicated above, and $w_{1t+1} = w_t, w_{2t} = w_t$ and $V_1 = V_2 = V_3$. Iterating on the Kalman filtering equations (2.38) and (2.35) with $\Sigma(t_0) = 0$, we obtain $\Sigma_t = 0,\ t \geq t_0,\ K_t = G,\ t \geq t_0$, and

$$(A - KC) = \begin{pmatrix} -c_1 & -c_2 & \dots & -c_{n-1} & -c_n \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix}.$$

It follows that

$$\widehat{E}_t x_{t+1} = \widehat{E}_t \begin{pmatrix} w_t \\ w_{t-1} \\ \vdots \\ w_{t-n+1} \end{pmatrix} = \begin{pmatrix} -c_1 & -c_2 & \dots & -c_{n-1} & -c_n \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix} \widehat{E}_{t-1} \begin{pmatrix} w_{t-1} \\ w_{t-2} \\ \vdots \\ w_{t-n} \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} y_t.$$

With $\Sigma_{t_0} = 0$, the above equation implies

$$\widehat{E}_t w_t = y_t - c_1 w_{t-1} - \dots - c_n w_{t-n}.$$

Thus the innovation w_t is recoverable from knowledge of y_t and n past innovations.

c. *Mixed moving average-autoregression*: We consider the univariate, mixed second-order autoregression, first-order moving average process

$$y_t = A_1 y_{t-1} + A_2 y_{t-2} + v_t + B_1 v_{t-1}$$

where v_t is a white noise with mean zero, $Ev_t^2 = V_1$ and $Ev_t y(s) = 0$ for $s < t$. The trick is getting this system into the state-space form is to define the state variables $x_{1t} = y_t - v_t$, and $x_{2t} = A_2 y_{t-1}$. With these definitions the system and measurement equations become

$$x_{t+1} = \begin{pmatrix} A_1 & 1 \\ A_2 & 0 \end{pmatrix} x(t) + \begin{pmatrix} B_1 + A_1 \\ A_2 \end{pmatrix} v_t \tag{2.48}$$

$$y_t = [1\ 0] x_t + v_t. \tag{2.49}$$

Notice that using (2.48) and (2.49) repeatedly, we have

$$\begin{aligned} y_t = x_{1t} + v_t &= A_1 x_{1t-1} + x_{2t-1} + (B_1 + A_1) v_{t-1} + v_t \\ &= A_1 (x_{1t-1} + v_{t-1}) + v_t + B_1 v_{t-1} + A_2 (x_{1t-2} + v_{t-2}) \\ &= A_1 y_{t-1} + A_2 y_{t-2} + v_t + B_1 v_{t-1} \end{aligned}$$

as desired. With the system and measurement equations (2.48) and (2.49), we have $V_1 = V_2 = V_3$,

$$A = \begin{pmatrix} A_1 & 1 \\ A_2 & 0 \end{pmatrix}, G = \begin{pmatrix} B_1 + A_1 \\ A_2 \end{pmatrix}, \ C = [1\ 0].$$

We start the system off with $\Sigma_{t_0} = 0$, so that the initial state is imagained to be known. With $\Sigma_{t_0} = 0$, recursions on (2.35) and (2.38) imply that $\Sigma_t = 0$ for $t \geq t_0$ and $K_t = G$ for $t \geq t_0$. Computing $A - KC$ we find

$$(A - KC) = \begin{pmatrix} -B_1 & 1 \\ 0 & 0 \end{pmatrix}$$

and we have

$$K = G = \begin{pmatrix} B_1 + A_1 \\ A_2 \end{pmatrix}.$$

Therefore the recursive prediction equations become

$$E_t y_{t+1} = [1\ 0] E_{t-1} x_t = E_{t-1} x_{1t}.$$

Recalling that $x_{2t} = A_2 y_{t-1}$, the preceding two equations imply that

$$E_t y_{t+1} = -B_1 E_{t-1} y_t + A_2 y_{t-1} + (B_1 + A_1) y_t. \tag{2.50}$$

Consider the special case in which $A_2 = 0$, so that the y_t obeys a first order moving average, first order autoregressive process. In this case (2.50) can be expressed

$$E_t y_{t+1} = B_1 (y_t - E_{t-1} y_t) + A_1 y_t,$$

which is a version of the Cagan-Friedman "error-learning" model. The solution of the above difference equation for $E_t y_{t+1}$ is given by the geometric distributed lag

$$\begin{aligned} E_t y_{t+1} = (B_1 + A_1) \sum_{j=0}^{m-1} (-B_1)^j y_{t-j} \\ + (-B_1)^m E_{t-m-1} y_{t-m}. \end{aligned}$$

For the more general case depicted in (c) with $A_2 \neq 0$, $E_t y_{t+1}$ can be expressed as a convolution of two geometric lag distributions in current and past y_t's.

d. Linear Regressions: Consider the standard linear regression model

$$y_t = z_t \beta + \varepsilon_t, \quad t = 1, 2, \ldots, T$$

where Z_t is a $1 \times n$ vector of independent variables, β is an $n \times 1$ vector of parameters, and ε_t is a random term with mean zero and variance $E\varepsilon_t^2 = \sigma^2$, and satisfying $E\varepsilon_t Z_s = 0$ for $t \geq s$. The least squares estimator of β based on t observations, denoted $\hat{\beta}_{t+1}$ is obtained as follows. Define the stacked vectors

$$Z_t = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_t \end{bmatrix}, \quad Y_t = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_t \end{bmatrix}.$$

Then the least squares estimator based on data through time t is given by

$$\hat{\beta}_{t+1} = (Z_t'Z_t)^{-1}Z_t'Y_t \tag{2.51}$$

with covariance matrix

$$E(\hat{\beta}_{t+1} - \bar{E}\hat{\beta}_{t+1})(\hat{\beta}_{t+1} - E\hat{\beta}_{t+1})' = \sigma^2(Z_t'Z_t)^{-1}. \tag{2.52}$$

For reference, we note that

$$\begin{aligned} \hat{\beta}_t &= (Z_{t-1}'Z_{t-1})^{-1}Z_{t-1}'Y_{t-1} \\ E(\hat{\beta}_t - E\hat{\beta}_t)(\hat{\beta}_t - E\hat{\beta}_t)' &= \sigma^2(Z_{t-1}'Z_{t-1})^{-1}. \end{aligned}$$

If $\hat{\beta}_t$ has been computed via (e), it is computationally inefficient to compute $\hat{\beta}_{t+1}$ via (2.51) when new data (y_t, z_t) arrive at time t. In particular, we can avoid inventing the matrix $(Z_t'Z_t)$ directly, by employing a recursive procedure for inverting it. This approach can be veiwed as an application of the Kalman filter. We explore this connection briefly.

We begin by noting how least squares estimators can be computed recursively via the Kalman filter. We let y_t in the Kalman filter be y_t in the regression model. We then set $x_t = \beta$ for all t, $V_{1t} = 0$, $V_{3t} = 0$, $V_{2t} = \sigma^2$, $w_{1t+1} = 0$, $w_{2t} = \varepsilon_t$, $A = I$, $C_t = z_t$. Let

$$\hat{\beta}_{t+1} = E\left[\beta \mid y_t, y_{t-1}, \dots y_1, z_t, z_{t-1}, \dots, z_1, \hat{\beta}_0\right],$$

where $\hat{\beta}_0$ is $\hat{x}_0$. Also, let $\Sigma_t = E(\hat{\beta}_t - E\hat{\beta}_t)(\hat{\beta}_t - E\hat{\beta}_t)'$. We start things off with a "prior" covariance matrix Σ_0. With these definitions, the recursive formulas (2.35) and (2.38) become

$$\begin{aligned} K_t &= \Sigma_t z_t'(\sigma^2 + z_t\Sigma_t z_t')^{-1} \\ \Sigma_{t+1} &= \Sigma_t - \Sigma_t z_t'(\sigma^2 + z_t\Sigma_t z_t')^{-1}z_t\Sigma_t \end{aligned} \tag{2.53}$$

Applying the formula $\hat{x}_{t+1} = (A - K_tC_t)\hat{x}_t + K_t y_t$ to the present problem with the above formula for K_t we have

$$\hat{\beta}_{t+1} = (I - K_t z_t)\hat{\beta}_t + K_t y_t. \tag{2.54}$$

We now show how (2.53) and (2.54) can be derived directly form (2.51) and (2.52). From a matrtix inversion formula (see Noble and Daniel [, p. 194]), we have that

$$(Z_t'Z_t)^{-1} = (Z_{t-1}'Z_{t-1})^{-1} - (Z_{t-1}'Z_{t-1})^{-1}z_t'(1 + z_t(Z_{t-1}'Z_{t-1}^1)^{-1}z_t')^{-1}z_t(Z_{t-1}'Z_{t-1})^{-1} \tag{2.55}$$

Multiplying both sides of (2.55) by σ^2 immediately gives (2.53). Use the right side of (2.55) to substitute for $(Z_t'Z_t)^{-1}$ in (2.51) and write

$$Z_t'Y_t = Z_{t-1}'Y_{t-1} + z_t'y_t$$

to obtain

$$\hat{\beta}_{t+1} = \frac{1}{\sigma^2}\{\Sigma_t - \Sigma_t z_t'(\sigma^2 + z_t\Sigma_t z_t')^{-1} z_t\Sigma_t\}$$

$$\cdot\{Z_{t-1}'Y_{t-1} + z_t'y_t\}$$

$$= \underbrace{\frac{1}{\sigma^2}\Sigma_t Z_{t-1}'Y_{t-1}}_{\hat{\beta}_t} - \underbrace{\Sigma_t z_t'(\sigma^2 + z_t\Sigma_t z_t')^{-1}}_{K_t}\underbrace{z_t}_{C_t}\underbrace{\frac{1}{\sigma^2}\Sigma_t Z_{t-1}'Y_{t-1}}_{\beta_t}$$

$$+ \underbrace{\Sigma_t Z_t'(\sigma^2 + z_t\Sigma_t Z_t')^{-1}}_{K_t} y_t$$

$$\hat{\beta}_{t+1} = (A - K_tC_t)\hat{\beta}_t + K_ty_t.$$

These formulas are evidently equivalent with those asserted above.

Computer Example: Using the Linear Regulator to Compute the Equilibrium of a Lucas-Prescott Model

This section reports the results of running the MATLAB program longluc4.m. This program computes the equilibrium of a linear quadratic version of Lucas and Prescott's model of investment under uncertainty. The program uses Lucas and Prescott's device of exploiting the fact that the rational expectations equilibrium of their model solves a fictitious social planning problem. For the linear quadratic version of their model (see, e.g., Sargent [1987, chapter XIV]), the social planning problem is a linear regulator problem. The program maps the social planning problem into a linear regulator. It uses a "doubling algorithm" to solve the problem.

You can edit this file and rerun the program in MATLAB to see how the equilibrium is sensitive to the specification of various demand and cost parameters. Here follows the output that appears on the screen in response to the command "longluc4".

```
longluc4
echo on
cla
```

This demonstration computes the solution of the social planning problem associated with a linear-quadratic version of Lucas and Prescott's 1971 model of investment under uncertainty. The model is altered to allow for a Romer externality.

There is a linear demand curve for output

$$p(t)=A(1) - A(2)^{*}Y(t) + u(t)$$

where p(t) is price, Y(t) is output and u(t) is a random shock to demand with an autoregressive process

$$u(t) = au(1)^{*}u(t-1) + \ldots + au(r)^{*}u(t-r) + eu(t)$$

where eu(t) is a white noise, and [au(1) . . . au(r)] is to be specified by the user.

The rental rate on capital w(t) also follows an rth order autoregression,

$$w(t) = aw(1)^{*}w(t-1) + \ldots + aw(r)^{*}w(t-r) + ew(t)$$

where ew(t) is a white noise, and [aw(1) . . . aw(r)] is to be specified.

```
pause      %press a key to continue the demonstration
cla
```

There are n identical firms. Each firm has production function

$$y(t) = f(1)^*k(t) + f(2)^*K(t)$$

where k(t) is capital of the representative firm and K(t)=n*k(t) is aggregate capital. We have Y(t)=n*y(t). Notice that aggregate output obeys

$$Y(t) = ff^*K(t)$$

where ff= f(1) + n*f(2). When f(2) is not zero, there is an externality.

```
pause      %press any key to continue the demonstration
cla
```

There is a fictitious social planner who chooses aggregate capital to maximize

$$\lim_{T\to\infty} \frac{1}{T} E \sum_{t=0}^{T} \{\text{consumer surplus(t) - producer surplus(t)}\}$$

where consumer surplus is given by

$$A(1)^*ff^*K(t) - (A(2)/2)^*(ff^*K(t))^2 + u(t)^*ff^*K(t)$$

and where producer surplus is given by

$$w(t)^*K(t) - (d/2n)^*(K(t)-K(t-1))^2$$

```
pause      %press a key to continue with the demonstration
cla
```

We'll set parameter values and then compute the equilibrium by mapping the social planning problem into a linear regulator.

```
pause      %press a key to start setting parameter values
cla
```

```
   A=[100 1],f=[1 .1],n=1

  A =

     100 1

  f =

     1.0000 0.1000

  n =

     1

   d=25

  d =

     25

pause     %press a key to set remaining parameters

cla

   au=[1.2 -.3]

  au =

     1.2000 -0.3000

   aw=[.9 0]

  aw =

     0.9000        0

pause     %press a key to continue

cla
```

We proceed to form the matrices (a,B,Q,R) for the linear regulator problem. The STATE vector is defined as x(t) = [K(t),1,u(t),u(t-1),w(t),w(t-1)]', and the CONTROL is defined as v(t)=(K(t)-K(t-1)). The transition matrix is called a and created as follows.

```
pause     %press a key to create the transition matrix a.

cla

  ff=f(1)+n*f(2);

  q=length(au);

  m=2*q+2;
```

```
a=zeros(m,m);
a(1,1)=1;
a(2,2)=1;
a(3:2+q,:)=[zeros(q,2),compn(au),zeros(q,q)];
a(3+q:m,:)  =[zeros(q,2),zeros(q,q),compn(aw)];
a
a =
```

1.0000	0	0	0	0	0
0	1.0000	0	0	0	0
0	0	1.2000	-0.3000	0	0
0	0	1.0000	0	0	0
0	0	0	0	0.9000	0
0	0	0	0	1.0000	0

```
pause      %Press a key to create B of the regulator.
B=zeros(m,1);
B(1,1)=1
B =

     1
     0
     0
     0
     0
     0

pause      %press a key to continue
cla
```

Now create R and Q of the regulator, where the regulator has the form

$$\max \lim_{T\to\infty} E\frac{1}{T}\sum_{t=0}^{\infty} \{ \text{x'*Q*x} + \text{v'*R*v} \}$$

subject to the law of motion

$$x(t+1) = a^*x(t) + B^*v(t) + \text{white noise}(t+1)$$

```
state='[K(t-1),1,u(t),u(t-1),w(t),w(t-1)]'
state =
[K(t-1),1,u(t),u(t-1),w(t),w(t-1)]
pause
cla
R=-d/(2*n);
Q=zeros(m,m);
Q(1,1)=-(ff^2)*A(2)/2
Q =
-0.6050  0  0  0  0  0
      0  0  0  0  0  0
      0  0  0  0  0  0
      0  0  0  0  0  0
      0  0  0  0  0  0
      0  0  0  0  0  0
Q(1,2)=A(1)*ff/2;
Q(2,1)=A(1)*ff/2;
Q(1,4)=ff/2;
Q(4,1)=ff/2;
Q(1,4+q)=-1/2;
Q(4+q,1)=-1/2;
pause     %press a key to give a and B
a
a =
1.0000       0  0  0  0  0
     0  1.0000  0  0  0  0
```

```
0   0   1.2000   -0.3000        0   0
0   0   1.0000         0        0   0
0   0        0         0   0.9000   0
0   0        0         0   1.0000   0

B
B =
1
0
0
0
0
0
pause      %press a key to give R and Q
Q
Q =
 -0.6050   55.0000   0   0.5500   0   -0.5000
 55.0000         0   0        0   0         0
       0         0   0        0   0         0
  0.5500         0   0        0   0         0
       0         0   0        0   0         0
 -0.5000         0   0        0   0         0

R
R =
-12.5000
pause      %Now solve the regulator problem.
F=double(a',B',Q,R);      %Working, please wait.
%       DONE.
pause      %Press a key to continue
```

The equilibrium control law for v(t) = K(t)-K(t-1) is given

$$K(t)-K(t-1) = -F^{*}x(t)$$

The state x(t) is given by state

```
state =
[K(t-1),1,u(t),u(t-1),w(t),w(t-1)]
pause     %Press a key to see the optimal value of F
F=F'
F =
0.1971 -17.9206 -0.1536 0.0370 0.1158 0
```

The optimal "closed loop" system is given by

$$x(t+1) = (a-B^{*}F) * x(t) + \text{white noise}(t+1)$$

```
pause     %press a key to see ABF = (a -B*F)
ABF=a-B*F
ABF =
```

0.8029	17.9206	0.1536	-0.0370	-0.1158	0
0	1.0000	0	0	0	0
0	0	1.2000	-0.3000	0	0
0	0	1.0000	0	0	0
0	0	0	0	0.9000	0
0	0	0	0	1.0000	0

```
state
state =
[K(t-1),1,u(t),u(t-1),w(t),w(t-1)]
pause     %press a key see eigenvalues of ABF
cla
eig(ABF)
ans =
```

```
0.8029

0.3551

0.8449

1.0000

0

0.9000

pause      %press a key to return to menu
```

This is the end of the output of "longluc4".

Computer Example: Using the Kalman Filter to Solve a Problem of Muth

This section reports the results of using the MATLAB program muthdem1.m. The program maps a classic signal extraction problem of Muth into the framework of the Kalman filter. The "doubling algorithm" is used to solve the matrix Ricatti equation that is associated with the Kalman filter.

The output response of the computer to the command "muthdem1" is now reproduced.

```
muthdem1

echo on

cla
```

This demonstration solves a signal extraction problem studied by Muth in order to rationalize "adaptive" expectations.

There is a hidden state variable x(t) that evolves according to an autoregressive process

$$x(t+1) = A * x(t) + e(t+1)$$

where A is a scalar (which Muth set equal to one) and e(t+1) is a white noise that is orthogonal to x(t). An agent observes a variable y(t), which is the sum of x(t) and a white noise:

$$y(t) = x(t) + u(t)$$

where E u(s)x(t)= 0 for all t and s. The variance of e(t+1) is given by Q and the variance of u(t) is given by R.

The problem is to find a (Wold) moving average representation for the observed variable y(t). We accomplish this by using the "Kalman filter".

```
pause     %Press a key to continue demonstration

cla
```

We use the Kalman filter to obtain an "innovations representation" of the form

$$xx(t+1) = A*xx(t) + K*a(t)$$

$$y(t) = xx(t) + a(t)$$

where xx(t) is E[x(t)—y(t),y(t-1), . . . , y(t(0)),xx(0)] a(t) is the one-step ahead prediction error in y(t), the so-called "innovation in y(t)", and K is the Kalman gain. From the innovations representation, which is a state space representation, we can obtain an a.r.m.a. representation for y(t) of the form

$$\mathrm{den}(L)y(t) = \mathrm{num}(L)a(t)$$

where den(L) and num(L) are scalar polynomials in the lag operator L.

```
pause      %press a key to continue
cla
```

You will be prompted for values of the parameters A, Q, and R.

NOTE: To obtain Muth's case, set A=1, so that the hidden signal follows a "random walk."

```
A=input('A= ')
  A=
A =
    1
C=1;
Q=input('give variance of state noise Q')
  give variance of state noise Q
Q =
    1
R=input('give variance of measurement noise R')
  give variance of measurement noise R
R =
    1
[K,s]=double(A,C,Q,R);
pause      %press a key to continue
cla
```

The value of the Kalman gain is given by

```
K

K =

    0.6180
```

The variance of the innovation a(t) in predicting y linearly from past values of y is given by

```
s

s =

    1.6180

pause      %Press a key to continue

cla
```

Now we'll give the a.r.m.a. representation for y(t)

$$\text{den}(L)y(t) = \text{num}(L)a(t)$$

Coefficients on L of power 0, 1, 2, ... [num,den]=ss2tf(A,K,C,1,1)

```
num =

    1.0000   -0.3820

den =

    1   -1

pause      %press a key to continue demonstration

cla
```

Muth showed that for a process of the form

$$(1 - L)y(t) = (1 - b\ L)a(t)$$

where a(t) is the innovation in y(t), the optimal one step ahead prediction of y(t+j) for j > 0 based on [y(t), y(t-1), . . .] is given by a geometric distributed lag

$$E[y(t+j) - y(t), y(t-1), \ldots] = (1-b)^* \sum_{k=0}^{\infty} \{b^{k*}\ y(t-k)\}$$

We invite you to experiment with this demonstration by varying Q and R while keeping A fixed (say at Muth's value of unity). In this way you can see the dependence of the parameters of the a.r.m.a. representation for y(t) on the ratio of Q to R.

```
pause       %press a key to return to menu
```

This terminates the output of "muthdem1". You can edit this file to solve signal extraction problems of your creation.

Computer Example: Using the Kalman Filter to Extract a Signal From a Signal Plus a Seasonal Noise

This section reports output from the MATLAB program recurseas.m. This program maps into the Kalman filter the problem of extracting the "signal" from the sum of a signal and a seasonal "noise". The doubling algorithm is used to solve the Ricatti equation associated with the Kalman filter.

The response to issuing the command "recurseas" is as follows.

```
recurseas
echo on
cla
```

USING THE KALMAN FILTER TO SEASONALLY ADJUST

NOTE: This demonstration takes several minutes, because relative to "classical" seasonal adjustment procedures, the ones used here substitute brute force and the Kalman filter for thought. If you have a train to catch, kill this demo by hitting "Ctrl,Break" and try another demo.

A reference for the techniques used here is Sargent's "Linear Control, Filtering, and Rational Expectations."

```
pause     %Press a key to proceed with demonstration.
cla
```

This program uses the Kalman filter to solve a "seasonal adjustment" problem that comes in the form of a signal extraction problem.

An observed process y(t) is the sum of three components:

a.) A "signal" f(t) that follows an autoregressive process

$$f(t) = a1(1)*f(t-1) + ... + a1(m)*y(t-m) + e1(t)$$

where e1(t) is a white noise with variance sig1.

b:) A "seasonal noise" s(t) that follows an a.r. process

$$s(t) = a2(1)*s(t\text{-}1) + \ldots + a2(r)*s(t\text{-}r) + e2(t)$$

where e2(t) is a white noise with variance sig2.

c.) A "measurement error" e3(t) which is a white noise with variance sig3.

NOTE: To approximate the case in which y(t) = f(t) + s(t), set sig3 equal to a very small positive number. The goal is to compute the linear least squares estimate

$$E[f(t) — y(t\text{-}1), y(t\text{-}2), \ldots].$$

```
pause       %Press a key to continue
cla
```

We solve the problem by mapping the system into state space notation, namely,

$$x(t+1) = A * x(t) + e(t)$$

$$y(t) = C * x(t) + v(t)$$

where x(t) is an (nx1) state vector and y(t) is a (kx1) vector of observations (in our example, k=1). The vector e(t) is an (nx1) vector white noise with covariance matrix Ee(t)e(t)'=Q. The vector v(t) is a (kx1) vector white noise which is orthogonal to e(s) for all t and s, and which has covariance matrix R.

For our example, the state vector x(t) will be given by

$$x(t) = [f(t)\ f(t\text{-}1)\ \ldots\ f(t\text{-}m)\ s(t)\ s(t\text{-}1)\ \ldots\ s(t\text{-}r)]'$$

while y(t) is simply the scalar observed variable.

```
pause       %Press a key to set parameters of a.r.  processes
cla
a1=[.9 0 0 0 0]
a1 =
```

0.9000 0 0 0 0

```
a2=[0 0 0 .9]
a2 =
```

0 0 0 0.9000

```
pause     %Press a key to form A matrix of state space representation
cla
A1=compn(a1);
A2=compn(a2);
[n,n1]=size(A1);[m,m1]=size(A2);
g1=zeros(n,m);
A=[A1,g1;g1',A2]
A =
  Columns 1 through 7
```

0.9000	0	0	0	0	0	0
1.0000	0	0	0	0	0	0
0	1.0000	0	0	0	0	0
0	0	1.0000	0	0	0	0
0	0	0	1.0000	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	1.0000	0
0	0	0	0	0	0	1.0000
0	0	0	0	0	0	0

```
  Columns 8 through 9
```

0	0
0	0
0	0
0	0
0	0
0	0.9000

```
    0         0
    0         0
1.0000        0

pause      %Press a key to form C.
cla
C=zeros(1,n+m);
C(1,1)=1;
C(1,n+1)=1
C =

    1   0   0   0   0   1   0   0   0

pause      %Press a key to set variance parameters and form R and Q
cla
Q=zeros(n+m);
Q(1,1)=1;
Q(n+1,n+1)=1
Q =

1   0   0   0   0   0   0   0   0
0   0   0   0   0   0   0   0   0
0   0   0   0   0   0   0   0   0
0   0   0   0   0   0   0   0   0
0   0   0   0   0   0   0   0   0
0   0   0   0   0   1   0   0   0
0   0   0   0   0   0   0   0   0
0   0   0   0   0   0   0   0   0
0   0   0   0   0   0   0   0   0

R=.0001;
pause      %Press a key to continue demonstration
cla
```

Now we'll use the Kalman filter to achieve the "innovations representation"

$$x(t+1) = (A - K*C)*x(t) + K*a(t)$$
$$y(t) = C*x(t) + a(t)$$

where a(t) = E[y(t) — y(t-1), y(t-2), ...], and K is the "Kalman gain". The process a(t) is the is the "innovation" in the y(t) process and has variance given by vara=C*S*C', where S is the covariance matrix of x(t+1) - x(t+1). The variable x(t+1) is the linear least squares projection

$$x(t+1) = E[x(t+1) — y(t), y(t-1), ...].$$

```
pause      %Press a key to compute K and S using the Kalman filter.
cla
[K,S]=double(A,C,Q,R);      %Working, please wait
pause      %Press a key to see Kalman gain K
K
K =
```

    0.4630
    0.5144
    0.1785
    0.0422
    -0.0442
    0.0397
    0.4856
    -0.1785
    -0.0422

```
pause      %Press a key to see state estimate covariance matrix
S
S =
  Columns 1 through 7
```

2.0743 1.1937 0.8407 0.7014 0.6891 -0.6202 -1.1936

1.1937	1.3263	0.9341	0.7794	0.7656	-0.6891	-1.3263
0.8407	0.9341	1.2362	0.9128	0.8017	-0.7215	-0.9341
0.7014	0.7794	0.9128	1.2312	0.9181	-0.8263	-0.7794
0.6891	0.7656	0.8017	0.9181	1.2257	-1.1031	-0.7657
-0.6202	-0.6891	-0.7215	-0.8263	-1.1031	1.9928	0.6891
-1.1936	-1.3263	-0.9341	-0.7794	-0.7657	0.6891	1.3263
-0.8407	-0.9341	-1.2362	-0.9128	-0.8016	0.7215	0.9341
-0.7014	-0.7794	-0.9128	-1.2312	-0.9181	0.8263	0.7794

```
Columns 8 through 9
```

-0.8407	-0.7014
-0.9341	-0.7794
-1.2362	-0.9128
-0.9128	-1.2312
-0.8016	-0.9181
0.7215	0.8263
0.9341	0.7794
1.2362	0.9128
0.9128	1.2312

```
pause      %Press a key to see variance of innovation to y.
vara=C*S*C'
vara =
    2.8268
pause      %Press a key to continue demonstration
cla
```

Notice that the innovations representation can be written

$$x(t+1) = A^*x(t) + K^*a(t)$$

$$y(t) = C^*x(t) + a(t)$$

This is equivalent with a Wold moving average representation for y(t), which we can represent in the rational form

$$\mathrm{den1}(L)y(t) = \mathrm{num1}(L)a(t)$$

```
pause      %Press a key to create num1 and den1.
[num1,den1]=ss2tf(A,K,C,1,1); pause      %Press a key to see num1
num1
num1 =
  Columns 1 through 7
    1.0000   -0.3973   -0.0737   -0.1265   -0.3184   0.0000   0
  Columns 8 through 10
       0      0      0
pause      %Press a key to see den1
den1
den1 =
    Columns 1 through 7
    1.0000   -0.9000   -0.0000   0.0000   -0.9000   0.8100   0
  Columns 8 through 10
       0      0      0
pause      %Press a key to continue demonstration
cla
```

We now calculate the spectrum of the filter num1(L)/den1(L).

```
pause      %Press a key to continue
st='spectrum of y(t) with unit variance of a(t)'
st =
spectrum of y(t) with unit variance of a(t)
sp=show(num1,den1,256,st);
```

See Figure 1

Spectrum of y(t) with unit variance of a(t)

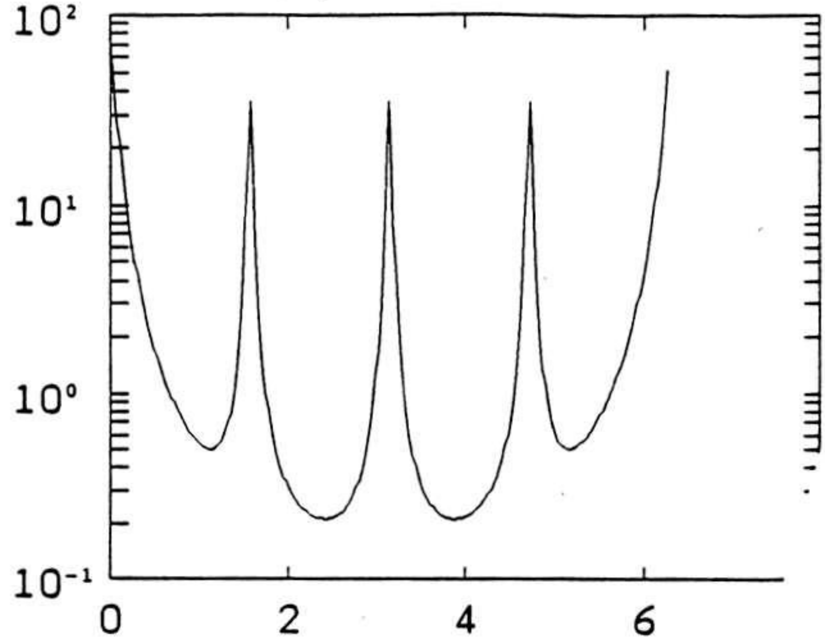

Figure 1

```
pause     %Press a key to continue
cla
```

Now we shall obtain the impulse response of x(t+1) to y(t). We rewrite the innovations representation as

$$x(t+1) = (A - K^*C)\ x(t) + Ky(t)$$

$$x(t) = eye^*x(t) + zeros\ ^*\ y(t)$$

We'll form the appropriate matrices, then use ss2tf to get a vector representation for x(t+1) of the form

$$den(L)x(t+1) = num(L)\ y(t)$$

where den(L) is a scalar polynomial in the lag operator and num(L) is a vector polynomial in the lag operator, with as many rows as components of x(t+1).

Notice that den(L) and num(L) contain all the information that we need to form each component of E[x(t+1)—y(t),...] Note that the first several components of E[x(t+1)—y(t),...] are E[f(t+1)—y(t),...], E[f(t)—y(t),...], E[f(t-1)—y(t),...], and so on. The lower rows of E[x(t+1)—y(t),...] thus correspond to finite two-sided seasonally adjusted series.

```
pause     %Press a key to form num and den
```

```
cla
C1= eye(n+m);D1=zeros(n+m,1);
[num,den]=ss2tf(A-K*C,K,C1,D1,1);% Working, please wait
pause     %Press a key to see num
num
num =
  Columns 1 through 7
```

0	0.4630	0.0000	0.0000	-0.0000	-0.4167	0
0	0.5144	-0.0000	-0.0000	0	-0.4630	-0.0000
0	0.1785	0.3537	-0.0000	0.0000	-0.1607	-0.3184
0	0.0422	0.1405	0.3537	-0.0000	-0.0380	-0.1265
0	-0.0442	0.0819	0.1405	0.3537	0.0397	-0.0737
0	0.0397	-0.0737	-0.1265	0.5816	-0.3933	0
0	0.4856	-0.3973	-0.0737	-0.1265	0.1446	0
0	-0.1785	0.6462	-0.3973	-0.0737	0.0342	0
0	-0.0422	-0.1405	0.6462	-0.3973	-0.0358	0

```
  Columns 8 through 10
```

0	0	0
0	0	0
-0.0000	0	0
-0.3184	0.0000	0
-0.1265	-0.3184	-0.0000
0	0	0
0	0	0
0	0	0
0	0	0

```
pause     %Press a key to see den
den
```

```
den =

  Columns 1 through 7

    1.0000   -0.3973   -0.0737   -0.1265   -0.3184   0.0000   0

  Columns 8 through 10

       0     0     0

pause      %Press a key to continue
cla
```

We now construct the impulse response of the first component of x(t+1) to an innovation in y(t). This is a representation for E[f(t+1)—y(t),...] in terms of a(t),a(t-1).... Notice from the innovations representation

$$x(t+1) = A^*x(t) + K^*a(t)$$

that the spectrum of the first component of x(t+1) is proportional to that of the first component of x(t+1). This follows from the whiteness of a(t).

```
pause      %Press a key to form the representation
num2=conv(num1,num(1,:));
den2=conv(den1,den);
pause      %Press a key to see num2
num2

num2 =

  Columns 1 through 7

       0   0.4630   -0.1839   -0.0341   -0.0586   -0.5641   0.1656

  Columns 8 through 14

    0.0307   0.0527   0.1327   -0.0000   0   0   0

  Columns 15 through 19

       0     0     0     0     0

pause      %Press a key to see den2
den2

den2 =
```

```
  Columns 1 through 7

    1.0000   -1.2973    0.2839   -0.0601   -1.1046    1.4542   -0.2555

  Columns 8 through 14

    0.0541    0.1841   -0.2579    0.0000         0         0         0

  Columns 15 through 19

         0         0         0         0         0

pause       %Press a key to plot spectrum of filter

st='spectrum of E[f(t+1)|y(t),y(t-1),...]'

st =

spectrum of E[f(t+1)|y(t),y(t-1),...]
```

See Figure 2

Spectrum of E[f(t+1)|y(t), y(t−1), . . .]

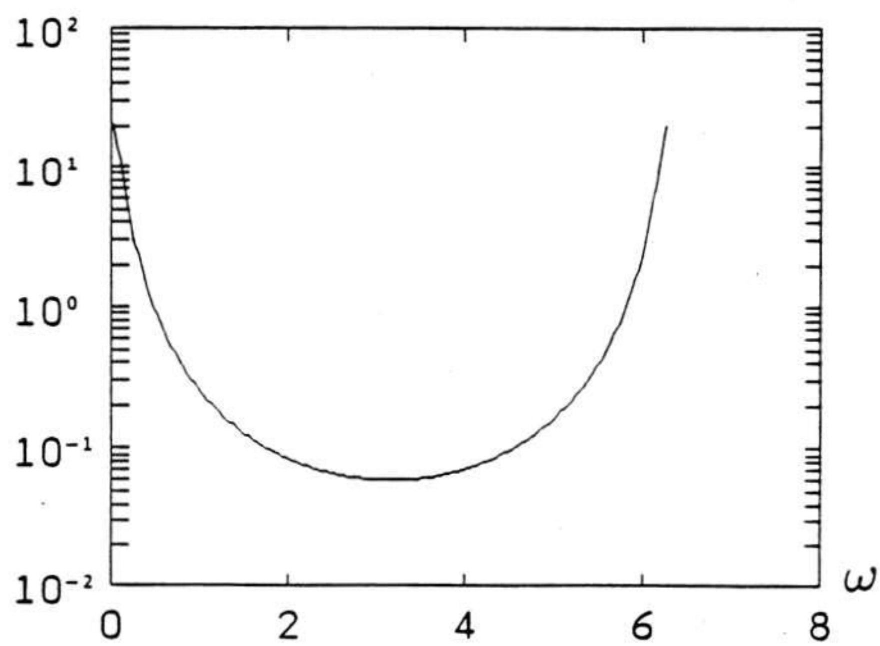

Figure 2

```
sp1=show(num2,den2,256,st);

pause       %Press key to continue

cla
```

We now construct the impulse response of the nth component of x(t+1) (i.e. E[f(t-n+1)—y(t), y(t-1),...]) to innovations a(t) in y(t). Notice that this is a finitely two sided signal extraction of f(t-n+1) based on past, present, and several future values of y(t). Because of the two-sidedness, there can occur "dips" in the spectral density of the seasonally adjusted process.

```
num3=conv(num1,num(n,:));
pause      %Press a key to plot spectrum
st='spectrum of E[f(t-n+1)|y(t),y(t-1),...]'
st =
spectrum of E[f(t-n+1)|y(t),y(t-1),...]
[sp3,ff]=show(num3,den2,256,st);
```

See Figure 3

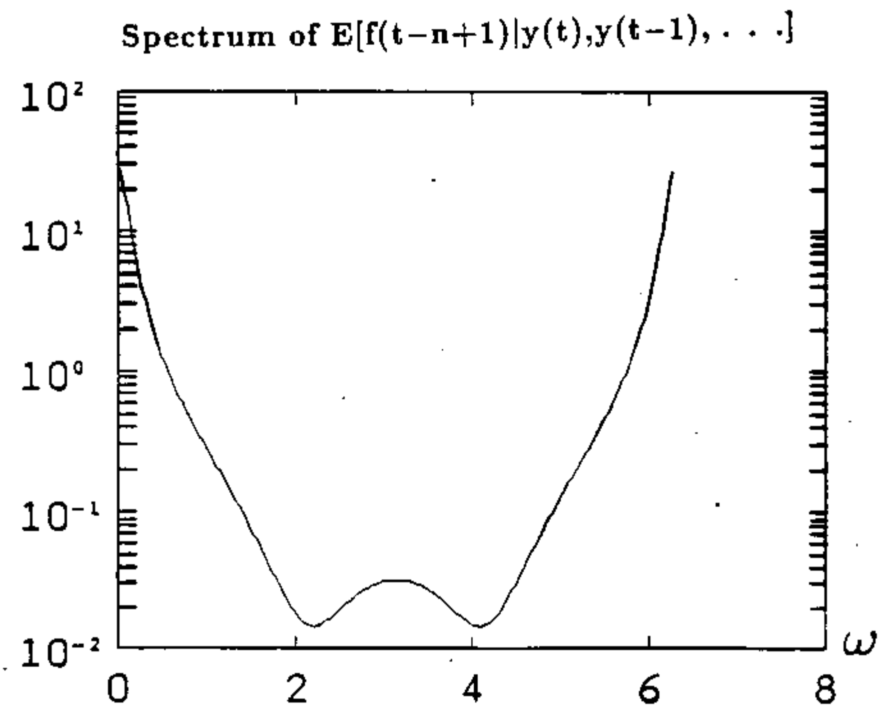

Figure 3

```
pause      %Press a key to continue
cla
mm=[sp1',sp3'];
st='spectra of one-sided and two-sided estimators of f(t)'
```

```
st =
spectra of one-sided and two-sided estimators of f(t)
semilogy(ff,mm),title(st),pause
```

See Figure 4

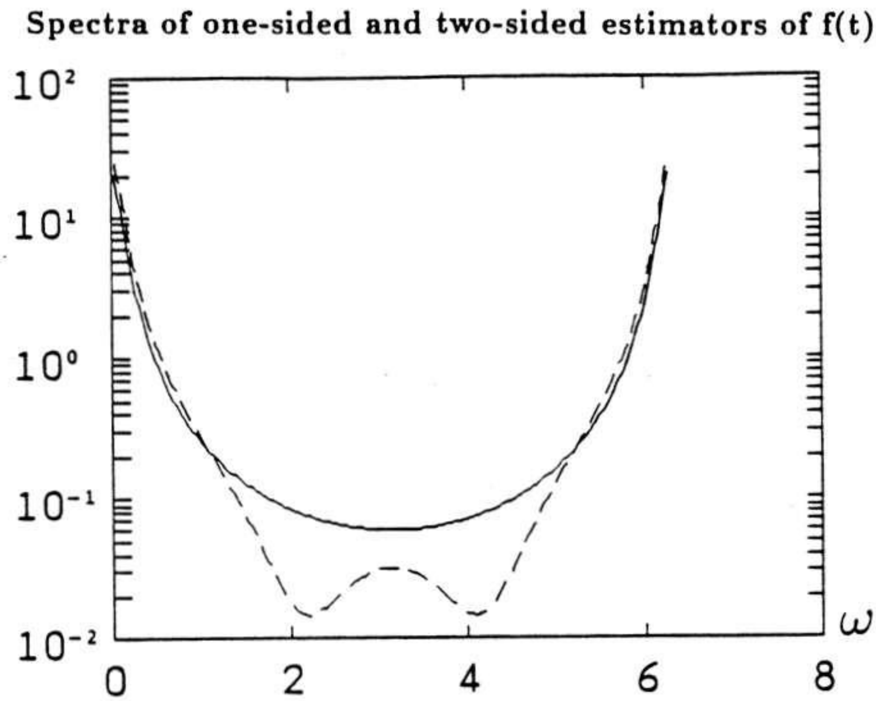

Figure 4

```
cla
```

Now we calculate the zeros of the numerator polynomial in the representation

$$\text{den2}(L)E[f(t-n+1)—y(t),\ldots] = \text{num3}(L)\ a(t)$$

```
r3=roots(num3); % Working, please wait
pause      %Press a key to see the roots
cla
r3
r3 =
    0
    0
```

```
0

0

3.4504

-0.7977 + 1.2983i

-0.7977 - 1.2983i

0.9740

0.9645

-0.9740

0.0000 + 0.9740i

0.0000 - 0.9740i

0.0350 + 0.7189i

0.0350 - 0.7189i

-0.6373

0.0001

-0.0000
```

```
pause     %Notice location of roots relative to one.  Press key
cla
```

NOTE: We have calculated the zeros of num3(z(-1)), which are the reciprocals of the zeros of num3(z). The "invertibility" condition is that the zeros of num3(z) be outside the unit circle,or that the zeros of num3(z(-1)) be inside the unit circle.

It is possible for some of these zeros to be outside the unit circle, reflecting the signal extraction version of the "invertibility problem" in rational expectations models discussed by Hansen and Sargent (1980)

```
pause     %Press a key to return to menu
```

This concludes the output of "recurseas". You can edit this file to create and solve your own signal extraction problems.

Appendix to Chapter 2

For reference we state the following theorems about linear least squares projections. We let Y be an $(n \times 1)$ vector of random variables and X be a $(h \times 1)$ vector of random variables. We assume that the following first and second moments exist:

$$EY = \mu_Y, \; EX = EX = \mu_X,$$

$$EXX' = S_{XX}, \; EYY' = S_{YY}, \; EYX' = S_{YX}.$$

Letting $x = X - EX$, $y = Y - EY$, we define the following covariance matrices

$$Exx' = \Sigma_{xx}, \; E'_{yy} = \Sigma_{yy}, \; Eyx' = \Sigma_{yx}.$$

We are concerned with estimating Y as a linear function of X. The estimator of Y which is a linear function of X and which minimizes the mean squared error between each component Y and its estimate is called the "linear projection of Y on X."

Definition A.1: The *linear projection* of Y on X is the affine function $\hat{Y} = AX + a_0$ which minimized $E\,trace\,\{(Y - \hat{Y})(Y - \hat{Y})'\}$ over all affine functions $a_0 + AX$ of X. We denote this linear projection as $\widehat{E}[Y \mid X]$, or sometimes as $\widehat{E}\,[Y \mid x, 1]$ to emphasize that a constant is included in the "information set".

The linear projection of Y on X, $\widehat{E}\,[Y \mid X]$ is also sometimes called the "wide sense expectation of Y conditional on X". We have

Theorem A.1:

$$E\,[Y \mid X] = \mu_y + \Sigma_{yx}\Sigma_{xx}^{-1}(X - \mu_x). \tag{A1}$$

Proof:

The theorem follows immediately by writing out $E\,trace\,(Y - \hat{Y})(Y - \hat{Y})'$, and completing the square, or else by writing out E trace $(Y - \hat{Y})(Y - \hat{Y})'$ and obtaining first-order necessary conditions ("normal equations") and solving them. ▮

Theorem A.2: (Orthogonality Principle):

$$E\left[(Y - \widehat{E}Y(x)) \mid X'\right] = 0$$

This states that the errors from the projection are orthogonal to each variable included in X.

Proof: Immediate from the normal equations. ▮

Theorem A.3:

(orthogonal regressions): Suppose that $X' = (X_1, X_2, \ldots, X_n)', \mu' = (\mu_{x1}, \ldots, \mu_{xn})'$ and that $E(x_i - \mu_{xj}) = 0$ for $i \neq j$. Then

$$(A2) \qquad \widehat{E}\left[Y \mid x_1, \ldots, x_n, 1\right] = \widehat{E}\left[Y \mid x_1\right] + \widehat{E}\left[Y \mid x_2\right] + \ldots + \widehat{E}\left[Y \mid x_n\right] - (n-1)\mu_y$$

Proof: Note that from the hypothesis of orthogonal regressors, the matrix Σ_{xx} is diagonal. Applying A1 then gives (A2). ▮

Chapter 3

Controllability and Stabilizability

1. Introduction

We shall eventually end up devoting most of our attention to optimal linear regulator problems that are time invariant, that is, problems for which the matrices R, Q, and W defining returns and the matrices A and B defining the transition law are all constant over time. For such time invariant problems, it will be of interest to have conditions that are sufficient to assure the following two outcomes: First, that iterates P_t produced by the matrix Ricatti difference equation converge; and second, when the matrix Ricatti difference equation does converge, that the optimal time invariant closed loop system $x_{t+1} = (A - BF)x_t$ is *stable*.

In this chapter and the next, we introduce the concepts of controllability and reconstructibility. It is in terms of these concepts that the desired convergence and stability theorems for the invariant linear regulator problem can be obtained. Roughly speaking, these concepts contribute to establishing stability of the optimal closed loop system in the following way. The optimal closed loop system $x_{t+1} = (A - BP)x_t$ will evidently be stable if it is both *desirable* and *feasible* to stablize the system through the application of feedback control. The concept of controllability and its specialization, the concept of stabilizability, tell whether or not A and B make it *feasible* to stabilize the system. The concepts of reconstructability and detectability describe whether R, Q, and W are such that it is *desirable* to stabilize the system. As we wade through the technical discussion of these concepts, it is useful to keep in sight how concepts will eventually be used to determine the stability of the system under the optimal control.

This chapter discusses the concepts of controllability and stabilizability. These concepts convey numerous insights into the structure of linear quadratic optimal control problems, as do the "dual" concepts of reconstructibility and detectability that are described in the following chapter.[1]

[1] The "Appendix to Chapters 3-4" lists a few theorems on linear algebra that will be used in the text. *Applied Linear Algebra*, second edition, by Ben Noble and James W. Daniel is one valuable reference on

2. Controllability

We consider the linear time invariant system

$$x_{t+1} = Ax_t + Bu_t \quad , \quad t \geq t_0 \tag{3.1}$$

where x_t is an $(n \times 1)$ vector of states, u_t a $(k \times 1)$ vector of controls, A an $(n \times n)$ matrix and B an $(n \times k)$ matrix. The matrices A and B are assumed to be independent of time. A solution of the first order difference equation (3.1) with a given initial vector $x_t = x_{t_0}$ at $t = t_0$ can be calculated recursively. In particular, notice that

$$x_{t+2} = Ax_{t+1} + Bu_{t+1} = A(Ax_t + Bu_t) + Bu_{t+1}$$

or

$$x_{t+2} = A^2 x_t + Bu_{t+1} + ABu_t .$$

Proceeding recursively to x_{t+j} gives

$$x_{t+j} = A^j x_t + Bu_{t+j-1} + ABu_{t+j-2} + \cdots + A^{j-1}Bu_t \quad , \quad j \geq 1 .$$

This can be written

$$x_{t+j} = A^j x_t + \sum_{i=1}^{j} A^{j-i} Bu_{t+i-1} \quad j \geq 1 \tag{3.2}$$

or equivalently as

$$x_t = A^{t-t_0} x_{t_0} + \sum_{s=t_0}^{t-1} A^{t-s-1} Bu_s, \; t = t_0 + 1, \; t_0 + 2, \; \ldots \tag{3.3}$$

It is useful to express (3.2) in the matrix form

$$x_{t+j} = A^j x_t + [B \;\; AB \;\; \cdots \;\; A^{j-1}B] \begin{bmatrix} u_{t+j-1} \\ u_{t+j-2} \\ \vdots \\ u_t \end{bmatrix} \tag{3.4}$$

Here the partitioned matrix $[B \;\; AB \;\; \cdots \;\; A^{j-1} B]$ is of dimension $n \times jk$ while the column vector $[u'_{t+j-1} \; u'_{t+j-2} \; \cdots \; u'_t]'$ is $jk \times 1$. Equation (3.4) reveals how the solution x_{t+j} to the

---

these and other theorems.

first order vector difference equation is the sum of $A^j x_t$, which represents the effects of the initial condition, and a linear combination of the columns of the matrix $[B \;\; AB \;\; \cdots \;\; A^{j-1}B]$, where the particular linear combination is determined by the vectors u_{t+j-1}, u_{t+j-2}, $\cdots$, u_t. To see this explicitly, let $(C)_i$ be the i^{th} column of a matrix C. Let the i^{th} element of the vector u_t be $u_{t,i}$. Then (3.4) can be written

$$
\begin{aligned}
x_{t+j} &= A^j x_t + (B)_1 u_{t+j-1,1} + (B)_2 u_{t+j-1,2} + \cdots + (B)_k u_{t+j-1,k} \\
&\quad + (AB)_1 u_{t+j-2,1} + (AB)_2 u_{t+j-2,2} + \cdots + (AB)_k u_{t+j-2,k} \cdots \\
&\quad + (A^{j-1}B)_1 u_{t,1} + (A^{j-1}B)_2 u_{t,2} + \cdots + (A^{j-1}B)_k u_{t,k}\,.
\end{aligned}
\tag{3.5}
$$

We now define the important concept of *complete controllability*.

Definition 3.1: The linear system $x_{t+1} = Ax_t + Bu_t$ is said to be *completely controllable* if the state of the system can be transferred from the zero state at any initial time t_0 to any terminal state $x_{t_1} = \bar{x}_1 \in R^n$ within finite time $(t_1 - t_0)$.

In other words, the system is completely controllable if and only if for any $\bar{x}_1 \in R^n$ there exists a $t_1 > t_0$ and a sequence u_{t_0}, u_{t_0+1}, $\cdots$, u_{t_1-1} such that starting from $x_{t_0} = 0$, the system moves to $\bar{x}_1$ at time t_1.

It is useful to remark that the definition implies that if the system is completely controllable then it can be moved from *any* initial state $\bar{x}_0$ at t_0 to any terminal state $\bar{x}_1$ at t_1 within finite time, $t_1 - t_0$. To verify this, let x_0 and x_1 be arbitrary points in R^n and suppose that it is desired to transfer the system from $\bar{x}_0$ at t_0 to $\bar{x}_1$ at some $t_1 > t_0$. If the system is completely controllable, it can be moved from the zero state to any state, in particular, to the state $\bar{x}_1 - A^{t_1-t_0}\bar{x}_0$ within finite time, $t_1 - t_0$. But from (3.1), the same sequence of inputs that moves the system from zero to $\bar{x}_1 - A^{t_1-t_0}\bar{x}_0$ at $t = t_1$ will also move the system from $\bar{x}_0$ to $\bar{x}_1$ at $t = t_1$.

At this point, we remind the reader of the Cayley-Hamilton theorem, which states that every square matrix satisfies its characteristic equation. That is, write the characteristic equation

$$| A - \lambda I | = 0$$

in the form

$$\phi_n \lambda^n + \phi_{n-1}\lambda^{n-1} + \cdots + \phi_0 = 0$$

where the ϕ_j's are scalar constants that depend on the elements of A. The Cayley-Hamilton theorem states that

$$\phi_n A^n + \phi_{n-1} A^{n-1} + \cdots + \phi_0 I = 0 .$$

Solving this equation for A^n gives

$$A^n = \sum_{j=0}^{n-1} g_j^n A^j \tag{3.6}$$

where the g_j^n's are constants that are functions of the ϕ_j's. Next notice that multiplying both sides of (3.6) by A gives

$$A^{n+1} = \sum_{j=0}^{n-1} g_j^n A^{j+1}$$

Using (3.6) to eliminate A^n from the right side of the above equation gives

$$A^{n+1} = \sum_{j=0}^{n-1} g_j^{n+1} A^j$$

where the g_j^{n+1} are again scalars. Continuing in the same fashion, it is established that

$$A^i = \sum_{j=0}^{n-1} g_j^i A^j \quad , \quad i \geq n \tag{3.7}$$

where the g_j^i are constants. Equation (3.7) expresses the i^{th} integer power of $A, i \geq n$, as a linear combination of the matrices $[I,\ A,\ \cdots,\ A^{n-1}]$. Thus, the columns of the matrices A^i, $i \geq n$, are linear combinations of the columns of $[I \ \ A \ \ \cdots \ \ A^{n-1}]$.

Post multiplying each side of (3.7) by B gives

$$A^i B = \sum_{j=0}^{n-1} g_j^i A^j B \quad , \quad i \geq n, \tag{3.8}$$

which shows that for $i \geq n$, the columns of $A^i B$ are linear combinations of the columns of the matrix $[B \ \ AB \ \ A^2B \ \ \cdots \ \ A^{n-1}B]$.

We are now in a position to state the following important theorem, which gives a characterization of complete controllability.

Theorem 3.1: The n-dimensional time invariant linear system $x_{t+1} = Ax_t + Bu_t$ is completely controllable if and only if the column vectors of the "controllability matrix"

$$P = (B \ \ AB \ \ A^2B \ \ \cdots \ \ A^{n-1}B)$$

span the n dimensional space that is, if and only if the rank of P equals n.

Proof: We first prove that complete controllability implies that the rank of P equals n. Repeating equation (3.2), the solution of the difference equation can be written

$$x_{t+j} = A^j x_t + \sum_{i=1}^{j} A^{j-i}\ B u_{t+i-1}, \quad j \geq 1.$$

Suppose $x_t = 0$, which as we saw was not restrictive when we discussed the definition of complete controllability. With x_t equal zero, the terminal state x_{t+j} is in the space spanned by the column vectors of the sequence of matrices $(B,\ AB,\ A^2B,\ \cdots)$. But it is an implication of (3.8) that this equals the space spanned by columns of the $n \times m \cdot k$ matrix $P = (B\ \ AB\ \ A^2B\ \ \cdots\ \ A^{n-1}B)$.

Thus, for all j, x_{t+j} is in the space spanned by the controllability matrix. If the columns of the controllability matrix do not span the n-dimensional space, then only states in the linear subspace spanned by P can be reached, which implies that the system is not completely controllable. This proves that if the system is completely controllable, then the rank of P equals n.

To prove the other direction of implication, suppose that the rank of P is n. Let us write the solution (3.4) with $j = n$ and $x_t = 0$,

$$x_{t+n} = [B\ \ AB\ \ \cdots\ \ A^{n-1}B] \begin{bmatrix} u_{t+n-1} \\ u_{t+n-2} \\ \vdots \\ u_t \end{bmatrix}$$

or

$$x_{t+n} = P \begin{bmatrix} u_{t+n-1} \\ u_{t+n-2} \\ \vdots \\ u_t \end{bmatrix}$$

We now set $x_{t+n} = x_1$, where x_1 is an arbitrary point in R^n, and we inquire whether there exists a vector $[u'_{t+n-1}, u'_{t+n-2}, \cdots, u'_t]'$ such that

$$x_1 = P \begin{bmatrix} u_{t+n-1} \\ u_{t+n-2} \\ \vdots \\ u_t \end{bmatrix}$$

or

$$x_1 = P\tilde{u},$$

where

$$\tilde{u} = [u'_{t+n-1}, \cdots, u'_t]'. \tag{3.9}$$

This question is equivalent with the question of whether the $n \times n \cdot k$ matrix P possesses an $nk \times n$ right inverse R which satisfies $PR = I_n$.

Notice that if $PR = I_n$, then (3.9) implies that $P\tilde{u} = x_1$, and $(PR)x_1 = P(Rx_1)$, so that $\tilde{u} = Rx_1$ is a solution of (3.9). This proves that existence of a right inverse of P implies the existence (but not the uniqueness) of a solution $\tilde{u}$ to (3.9) for every $x_1 \in \mathrm{R}^n$. From a theorem in linear algebra (Noble and Daniel [, p. 97]) P has a right inverse if and only if the rank of P equals n. Also, if the rank of P is n, then PP' is nonsingular (Noble and Daniel [ ]). From this fact, it can be directly verified that one right inverse of P is $R = P^T(PP^T)^{-1}$. Thus, if the rank of P equal n, there exists at least one sequence of controls $[u'_{t+n-1}, u'_{t+n-2}, \cdots, u'_t]'$ which drives the system from zero to x_1 in n time periods. Given a right inverse R of P, such a sequence of controls can be computed from

$$\begin{bmatrix} u_{1t+n-1} \\ u_{1t+n-2} \\ \vdots \\ u_t \end{bmatrix} = (Rx_1) \tag{3.10}$$

This proves that if the rank of P is n, then the system is completely controllable. This completes the proof of the theorem. ▮

We remark that more has been proved then was stated in the theorem. In particular, we have proved that if the system is completely controllable, then (3.9) or (3.10) implies that it is possible to move the system from any initial state x_0 at t_0 to any other state x_1 within at most n periods.

The following definition will prove useful:

Definition 3.2: The *controllable subspace* of the linear time invariant system $x_{t+1} = Ax_t + Bu_t$ is the linear subspace consisting of the states that can be reached from the zero state in finite time.

We immediately have the following theorem:

Theorem 3.2: The controllable subspace of the n-dimensional linear time invariant system $x_{t+1} = Ax_t + Bu_t$ is the linear subspace spanned by the columns of the controllability matrix $P = \{B \ \ AB \ \ \cdots \ \ A^{n-1}B\}$.

Proof: The proof of this theorem is contained in the proof of theorem 3.1.

We shall use the following theorem.

Theorem 3.3: The controllable subspace of the system $x_{t+1} = Ax_t + Bu_t$ is *invariant under* A; that is, if x is in the controllable subspace, Ax is also in the controllable subspace.

Proof: Let the controllable subspace be denoted by $C = R[B\vdots AB\vdots \cdots \vdots A^{n-1}B]$ where $R(D)$ denotes the range space of the matrix D. If x is an element of the space C, then x is in the space spanned by the column vectors of the controllability matrix P. Notice that if x belongs to C, then Ax is in the linear subspace spanned by column vectors of $[AB \ \ A^2B \ \ \cdots \ \ A^nB]$. Equation (3.8) implies that the column vectors of A^nB depend linearly on the column vectors of P. Therefore, Ax is an element of C. ▮

Heuristically, notice that if x belongs to C, we can drive the state from zero to x in at most n periods. Having arrived at x at period t_1, we can get to Ax in period $t_1 + 1$ by setting $u_{t_1} = 0$. However, we can get to Ax directly from zero faster, that is in at most n steps, as the proof indicates.

We also have the following useful theorem.

Theorem 3.4: An initial state x_0 belonging to C at time t can be transferred to any terminal state x_1 in C in at most n periods.

Proof: Repeating the solution (3.2) for $j = n$ gives

$$x_{t+n} = A^n x_t + \sum_{i=1}^{n} A^{n-i} B u_{t+i-1}.$$

Now if x_t belongs to C then $A^n x_t$ belongs to C by theorem 3.3. Further, the above equation shows that any input sequence $[u_t, u_{t+1}, \ldots, u_{t+n-1}]$ that transfers the zero state to $x_1 - A^n x_0$ also transfers x_0 to x_1. Such an input sequence exists since $x_1 - A^n x_0$ is in the controllable subspace C. ▮

3. The Controllability Canonical Form

We now proceed to the construction of the *controllability canonical form*. This form will be especially important for the class of economic models that we shall consider, since many of them are naturally specified to be in controllability canonical form.

We consider the n dimensional linear time invariant system $x_{t+1} = Ax_t + Bu_t$. Let the rank of the controllability matrix $P = [B, AB, \ldots, A^{n-1}B]$ equal $m \leq n$. So the dimension of the controllable subspace C equals $m \leq n$. Choose any basis for C consisting of the n x 1 column vectors $e_1, e_2, \ldots, e_m$. Let $e_{m+1}, e_{m+2}, \ldots, e_n$ be $(n-m)$ linearly independent n x 1 vectors which together with $e_1, \ldots, e_m$ span the entire n-dimensional space. Form

$$T = (T_1,\ T_2)$$

$$\text{where} \quad T_1 = (e_1,\ \cdots,\ e_m) \quad .$$

$$T_2 = (e_{m+1},\ \cdots,\ e_n)$$

Now introduce the transformed state vector

$$Tx_t' = x_t \ . \tag{3.11}$$

Substituting (3.11) into the state difference equation gives

$$Tx_{t+1}' = ATx_t' + Bu_t$$

or

$$x_{t+1}' = T^{-1}ATx_t' + T^{-1}Bu_t \ . \tag{3.12}$$

Now partition T^{-1} as

$$T^{-1} = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix}$$

where U_1 is $m \times n$ and U_2 is an $(n-m) \times n$ matrix. Then we have

$$T^{-1}T = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} (T_1 T_2) = \begin{bmatrix} U_1T_1 & U_1T_2 \\ U_2T_1 & U_2T_2 \end{bmatrix} = \begin{bmatrix} I_m & 0 \\ 0 & I_{n-m} \end{bmatrix}$$

This implies that $U_2T_1 = 0$. Recall that T_1 is composed of vectors $e_1, \ \ldots, \ e_m$ that span the controllable subspace. Then the equality $U_2T_1 = 0$ implies that $U_2x = 0$ for any vector x belonging to the controllable subspace C. That is, if x belongs to C then it can be written as $x = T_1 y$ for some $m \times 1$ vector y. So we have $U_2 x = U_2 T_1 y = 0$, as an implication of $U_2 T_1 = 0$. Thus we have

$$U_2 x = 0 \tag{3.13}$$

for any x belonging to C.

With the preceding partitioning of T and U, we can write

$$T^{-1}AT = \begin{bmatrix} U_1 A T_1 & U_1 A T_2 \\ U_2 A T_1 & U_2 A T_2 \end{bmatrix}$$

$$T^{-1}B = \begin{bmatrix} U_1 B \\ U_2 B \end{bmatrix}.$$

By construction, all columns of T_1 are in the controllable subspace C. Since the controllable subspace is invariant under A by theorem3.3, all columns of AT_1 are also in C. It then follows from (3.13) that

$$U_2 A T_1 = 0.$$

The columns of B are obviously in the controllable subspace, since B belongs to P. Therefore, we also have

$$U_2 B = 0.$$

Thus, we have established that (3.12) assumes the form

$$\begin{bmatrix} x_{1t+1} \\ x_{2t+1} \end{bmatrix} = \begin{bmatrix} A'_{11} & A'_{12} \\ 0 & A'_{22} \end{bmatrix} \begin{bmatrix} x'_{1t} \\ x'_{2t} \end{bmatrix} + \begin{bmatrix} B'_1 \\ 0 \end{bmatrix} u_t \tag{3.14}$$

or

$$x'_{t+1} = A' x'_t + B' u_t$$

where $A'_{11} = U_1 A T_1$, $A'_{12} = U_1 A T_2$, $A'_{22} = U_2 A T_2$, $B'_1 = U_1 B$, $A' = T^{-1}AT, B' = T^{-1}B, x'_{1t}$ is an $(m \times 1)$ vector and x'_{2t} is an $(n-m) \times 1$ vector.

Equation (3.14) is called the controllability canonical form of the linear system $x_{t+1} = Ax_t + Bu_t$. The importance of this canonical form is partly due to the properties exhibited in the following theorem:

Theorem 3.5: In the controllability canonical form (3.14), the pair (A'_{11}, B'_{11}) is controllable.

Proof: It is sufficient to establish that the rank of the matrix

$$P'_{11} = [B'_1, \ A'_{11}B'_1, \ ..., \ A'^{n-1}_{11}B'_1]$$

equals m. We have

$$\begin{aligned} P'_{11} &= [U_1B, \ U_1AT_1U_1B, \ \ldots, \ (U_1AT_1)^{n-1}U_1B] \\ &= U_1[B, \ AB, \ \ldots, \ A^{n-1}B] \\ &= U_1P. \end{aligned}$$

The $(m \times n)$ matrix U_1 has rank m by construction, and the $(n \times nk)$ matrix P has rank m by assumption. Since the columns of T_1 form a basis for the range space of P, we have rank $(P'_{11}) = rank(U_1P) = rank(U_1T_1) = rank(I_m) = m$. This proves the theorem. ∎

It is useful to note that the controllability matrix for the pair $(A', \ B')$ of the controllability canonical form is

$$\begin{aligned} T^{-1} = P' &= [B', \ A'B', \ \ldots, \ A'^{n-1}B'] \\ &= \begin{bmatrix} B'_1, & A'_{11}B'_1 & & A'^{n-1}B' \\ 0 & 0 & & 0 \end{bmatrix} \end{aligned}$$

An alternative proof of theorem 3.5 notes that P has ranks m, and that T^{-1} is nonsingular, which imply that the rank of P' is m. This in turn implies that the matrix P'_{11} has rank m.

At the cost of being redundant, we find it useful to summarize the result of theorem 3.5 and the discussion leading up to it in the following theorem.

Theorem 3.6: Consider the linear time invariant system $x_{t+1} = Ax_t + Bu_t$. There exists a nonsingular transformation matrix T such that the transformed state $x'_t = T^{-1}x_t$ is in the *controllability canonical form*

$$\begin{bmatrix} x'_{1t+1} \\ x'_{2t+1} \end{bmatrix} = \begin{bmatrix} A'_{11} & A'_{12} \\ 0 & A'_{22} \end{bmatrix} \begin{bmatrix} x'_{1t} \\ x'_{2t} \end{bmatrix} + \begin{bmatrix} B'_1 \\ 0 \end{bmatrix} u_t$$

where $A' = T^{-1}AT$ and $B' = T^{-1}B$, x'_1 is $m \times 1$ and x'_2 is $(n - m) \times 1$, where $m = rank[B, \ AB, \ \ldots, \ A^{n-1}B]$, and A'_{11} is $m \times m$ and B'_1 is $(m \times 1)$. The pair $(A'_{11}, B'_1$ is controllable.

From the method of constructing the controllability canonical form, it is evident that it is not unique. This is true because T_1 can be chosen as any m $(n \times 1)$ column vectors that form a basis for the controllable subspace C. It follows also that T_2 is not unique. However, it can be proved that for any of the controllability canonical forms (3.14) produced by selecting different admissible T_1 and T_2 matrices, the eigenvalues of A'_{11} are the same regardless of the choice of T_1 and T_2, and that the eigenvalues of A'_{22} are also independent of the choice of particular admissible T_1 and T_2.

We state these facts in the form of the following theorem:

Theorem 3.7: Let T be a nonsingular matrix $(T_1 \quad T_2)$ where the columns of T_1 form a basis for the controllable subspace of a pair $(A, \ B)$, and let $(T_1 \quad T_2)$ form a basis for R^n. Let $\bar{T} = (\bar{T}_1 \quad \bar{T}_2)$ be another nonsingular $(n \times n)$ matrix whose first m columns span the controllable subspace of the same pair $(A, \ B)$. Then consider

$$A' = \begin{bmatrix} A'_{11} & A'_{12} \\ 0 & A'_{22} \end{bmatrix} = T^{-1}AT$$

and

$$\bar{A}' = \begin{bmatrix} \bar{A}_{11} & \bar{A}_{12} \\ 0 & \bar{A}_{22} \end{bmatrix} = \bar{T}^{-1}A\bar{T} \, .$$

The eigenvalues of A'_{11} equal those of $\bar{A}_{11}$, and the eigenvalues of A'_{22} equal those of $\bar{A}_{22}$.

Proof: From $A' = T^{-1}AT$ and $\bar{A} = \bar{T}^{-1}A\bar{T}$, where T and $\bar{T}$ are nonsingular, we have

$$A' = (T^{-1}\bar{T})\bar{A}(\bar{T}^{-1}T) = (T^{-1}\bar{T})\bar{A}(T^{-1}\bar{T})^{-1} \tag{3.17}$$

where $(T^{-1}\bar{T})$ is a nonsingular matrix.

Letting $T^{-1} = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix}$, we have

$$T^{-1}\bar{T} = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} (\bar{T}_1 \ \bar{T}_2) = \begin{bmatrix} U_1\bar{T}_1 & U_1\bar{T}_2 \\ U_2\bar{T}_1 & U_2\bar{T}_2 \end{bmatrix}$$

Using the above established fact that $x \in C$ implies $U_2x = 0$, we have $U_2\bar{T}_1 = 0$ because $\bar{T}_1$ is a basis for C. Therefore, we have

$$T^{-1}\bar{T} = \begin{bmatrix} U_1\bar{T}_1 & U_1\bar{T}_2 \\ 0 & U_2\bar{T}_2 \end{bmatrix} . \tag{3.18}$$

Similarly, we have

$$\bar{T}^{-1}T = \begin{bmatrix} \bar{U}_1T_1 & \bar{U}_1T_2 \\ 0 & \bar{U}_2T_2 \end{bmatrix}$$

Calculating A' using (3.17) and (3.18) gives

$$\begin{bmatrix} A'_{11} & A'_{12} \\ 0 & A'_{22} \end{bmatrix} = \begin{bmatrix} U_1\tilde{T}_1A_{11}\tilde{U}_1T_1 & (U_1\tilde{T}_1A_{11}U_1T_2 + U_1T_1A_{12}\tilde{U}_2T_2 & + U_1\tilde{T}_2A_{22}\tilde{U}_2T_2) \\ & 0 & U_2\tilde{T}_2\tilde{A}_{22}\tilde{U}_2T_2 \end{bmatrix}$$

Thus, we have

$$A'_{11} = U_1\tilde{T}_1\tilde{A}_{11}\tilde{U}_1T_1$$

$$A'_{22} = U_2\tilde{T}_2\tilde{A}_{22}\tilde{U}_2T_2$$

Upon noting that $(U_1\tilde{T}_1)^{-1} = \tilde{U}_1T_1$ and $(U_2\tilde{T}_2)^{-1} = \tilde{U}_2T_2$, the above equations imply that A'_{11} and $\tilde{A}_{11}$ are related by a similarity transformation, and that A'_{22} and $\tilde{A}_{22}$ are related by a similarity transformation. Therefore, the eigenvalues of A'_{11} equal those of $\tilde{A}_{11}$, and the eigenvalues of A'_{22} equal those of $\tilde{A}_{22}$. ∎

The preceding discussion motivates the following definitions:

Definition 3.3 The characteristic values of A'_{11} are called the *controllable poles* of the system $(A,\ B)$.

Definition 3.4: The characteristic values of A'_{22} are called the *uncontrollable poles* of the system $(A,\ B)$.

4. Stabilizability

Now suppose that A has n distinct eigenvalues. Recall the eigenvalue decomposition of A,

$$A = S\ \Lambda\ S^{-1}$$

where $\Lambda = diag(\lambda_1,\ \ldots,\ \lambda_n)$ is the diagonal matrix whose entries are eigenvalues of A, and $S = (s_1,\ \ldots,\ s_n)$ is the matrix whose columns are eigenvectors of A. Let us represent S^{-1} as

$$S^{-1} = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{bmatrix}$$

where the f_j are $(1 \times n)$ matrices. Consider the homogeneous linear time invariant difference equation

$$x_{t+1} = Ax_t$$

whose solution is

$$x_{t+j} = A^j x_t.$$

Using the eigenvalue decomposition of A, this solution can be represented

$$x_{t+j} = S\ \Lambda^j S^{-1} x_t$$

or

$$x_{t+j} = \sum_{i=1}^{n} \lambda_i^j s_i f_i x_t. \tag{3.19}$$

Equation (3.19) shows how the behavior of the homogeneous system depends on the eigenvalues of A.

We now make the following definition:

Definition 3.5: The homogeneous time invariant linear system $x_{t+1} = Ax_t$ is said to be *stable* if for any x_{t_0} belonging to $\mathrm{R}^n, \lim_{j\to\infty} x_{t+j} = 0$.

From equation (3.19), the following theorem is immediate:

Theorem 3.8: The homogeneous time invariant linear system is stable if and only if the eigenvalues of A are strictly less than unity in modulus.

If the eigenvalues of A are strictly less than unity in modulus, we also speak of A as a *stable matrix.*

We now indicate that theorem 3.8 continues to hold in the case that the eigenvalues of A are not all distinct. If the eigenvalues of A are not distinct, equation (3.19) does not hold, but a suitable generalization of it does. We recall several facts from linear algebra. Let the $(m \times n)$ matrix A have $k \leq n$ distinct eigenvalues, $\lambda_1, \ldots, \lambda_k$. Let m_i be the multiplicity of the eigenvalue λ_i. Associated with each eigenvalue λ_i there can be anywhere between one and m_i linearly independent eigenvectors.

Define the matrices

$$M_i = (A - \lambda_i I)m_i$$

and let

$$N_i = \mathcal{N}(M_i)$$

be the null space of M_i. Then it follows that (a) the dimension of the linear subspace N_i is $m_i, i = 1, \ldots, k_i$ and (b) each vector x in R^n can be expressed uniquely as a sum $x = x_1 + x_2 + \ldots + x_k$ where x_i belongs to N_i. (See Kwakernaak and Sivan [p. 19]). The x_i can be expressed as linear combinations of the eigenvectors and "generalized eigenvectors" corresponding to λ_i, which we proceed to define and describe how to compute.

We first define a *Jordan block* matrix J_i as a square matrix whose elements are zero except for those on the principal diagonal, which are all equal to unity, and those in the first superdiagonal, which all equal unity. Thus,

$$J_i = \begin{bmatrix} \lambda_i & 1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda_i & 1 & 0 & \ldots & 0 \\ \cdot & & & & & \\ \cdot & & & & & \\ \cdot & \cdot & \cdot & \cdot & \ldots & \lambda_i \end{bmatrix}$$

The number λ_i is taken to be an eigenvalue of A.

We now state the following theorem.

Theorem 3.9: Let A be a square matrix. Then there exists a nonsingular transformation matrix T which can be partitioned $T = (T_1,\ T_2, \ldots,\ T_k)$, where T_i has m_i columns, such that

$$A = TJT^{-1}$$

where J is block diagonal and is composed of Jordan blocks along the diagonal. In particular, associated with each linearly independent eigenvector of A there is one Jordan block in J, with its associated eigenvalue. For each eigenvalue of A there are as many Jordan blocks as there are linearly independent eigenvectors associated with it. The vectors in T_i form a basis for N_i and are either eigenvectors of "generalized eigenvectors." (If an eigenvalue λ_i has multiplicity $m_i > 1$ and there is only one linearly independent eigenvalue associated with it, that eigenvector can be taken as the first column of T_i. In this case, the Jordan block corresponding to λ_i has dimension $m_i \times m_i$). More generally, the matrix J can be partitioned as

$$J = \begin{bmatrix} J_1 & & 0 \\ & \ddots & \\ 0 & & J_k \end{bmatrix}$$

where each block J_i has dimension $m_i \times m_i$. Each $(m_i \times m_i)$ block J_i is of the form

$$J_i = \begin{bmatrix} J_{i1} & & 0 \\ & \ddots & \\ 0 & & J_{i\ell_i} \end{bmatrix}$$

where each J_{ij} is a Jordan block associated with eigenvalue λ_i, and ℓ_i is the number of linearly independent eigenvectors corresponding to λ_i. This completes the statement of the theorem. (See Kwakernaak and Sivan [ ] or Nobel and Daniel [ ]).

From the equation $AT = TJ$, Noble and Daniel describe the following method of computing the columns of T. Let these columns of T be $v_1, \ldots, v_n$. Then from the form of J and the equation $AT = TJ$ if follows that

$$Av_i = \lambda v_i + \gamma_i v_{i-1}$$

where λ_i is either 0 or 1 depending on J and where λ is a characteristic value of A. Partition the block T_i of T corresponding to the subpartitioning of J_i as $(T_{i1}, \ldots, T_{i\ell_i})$. Then γ_i is zero whenever the corresponding column v_i of T is the first column of a subblock. If $\gamma_i = 0, v_i$ is an eigenvector of A corresponding to λ. Thus the first column of each subblock T_{ij} can be taken as an eigenvector corresponding to λ, while the remaining columns follow recursively from the above equation with $\gamma_i = 1$. The remaining columns of T_{ij} generated in this way are called "generalized eigenvectors" of A.

It is useful to compute integer powers of an $r \times r$ Jordan block matrix J_k. Letting

$$J_k = \begin{bmatrix} \lambda & 1 & 0 & \ldots & 0 \\ 0 & \lambda & 1 & 0 & \ldots \\ \vdots & & & & \\ 0 & . & . & . & \lambda \end{bmatrix},$$

it is readily verified that

$$J_k^t = \begin{bmatrix} \lambda^t & \binom{t}{1}\lambda^{t-1} & \binom{t}{2}\lambda^{t-2} & \ldots & \binom{t}{r}\lambda^{t-r} \\ 0 & \lambda^t & \binom{t}{1}\lambda^{t-1} & \ldots & \binom{t}{r-1}\lambda^{t-r+1} \\ \vdots & & & & \\ 0 & 0 & 0 & \ldots & \lambda^t \end{bmatrix}$$

From the representation $A = TJT^{-1}$ we have that $A^t = TJ^tT^{-1}$. In the case where generalized eigenvectors are included in T, it follows that the solution $x_t = A^t x_0$ has components that behave as $\binom{t}{r}\lambda^t$. Since $\binom{t}{r}\lambda^t$ goes to zero as t goes to infinity if and only if $|\lambda|< 1$, theorem 3.8 about stable matrices also holds for the case in which the eigenvalues are not distinct.

In summary, associated with each eigenvalue λ_i of multiplicity of m_i the $n \times n$ matrix A there is associated a set of m_i linearly independent eigenvectors and generalized eigenvectors that span the null space $\mathcal{N}(A - \lambda_i I)m_i$.

In the text below, we shall on several occasions state and prove theorems about the spaces spanned by the eigenvectors corresponding to particular collections of eigenvalues. For simplicity, in our proofs we shall assume that the eigenvalues are distinct, and so use the eigenvalue decomposition $A = S\Lambda S^{-1}$. However, the argument in each of the proofs goes through in the case of repeated eigenvalues if we use the Jordan decomposition $A = TJT^{-1}$ and interpret the "space spanned by the eigenvectors corresponding to λ_i" to mean the "space spanned by the eigenvectors and generalized eigenvectors corresponding to λ_i."

The following definition will prove very useful:

Definition 3.6: Consider the n-dimensional linear time invariant system $x_{t+1} = Ax_t$. Suppose that A has n distinct eigenvalues. We define the *stable subspace* of this system as the real linear subspace spanned by those eigenvectors of A that correspond to eigenvalues with moduli strictly less than unity. The *unstable subspace* of the system is the real subspace spanned by those characteristic vectors that correspond to eigenvalues with moduli greater than or equal to unity.

We note that as a consequence of this definition and of the eigenvalue decomposition $A = S\Lambda S^{-1}$, it follows that any vector x_t in R^n can be represented uniquely as

$$x_t = x_{st} + x_{ut}$$

where x_{st} is in the stable subspace of A and x_{ut} is in the unstable subspace of A.

The following concept is very useful because it is instrumental in characterizing a set of conditions that are sufficient to guarantee both convergence of the matrix Riccati equation and stability of the closed loop system.

Definition 3.8: The linear time invariant system $x_{t+1} = Ax_t + Bu_t$ is said to be *stabilizable* if its unstable subspace is contained in its controllable subspace. That is, the system $(A,\ B)$

is stabilizable if x the condition that belongs to the unstable subspace implies that x belongs to the controllable subspace.

The following two theorems are immediate:

Theorem 3.10: Any stable time invariant system is stabilizable.

Proof: The unstable subspace is empty. ▮

Theorem 3.11: Any controllable system is stabilizable.

Proof: The controllable subspace is R^n. ▮

The property displayed in the following lemma is useful:

Lemma 3.1: The controllable subspace of the pair (A', B') of the controllability canonical form (3.14) is spanned by the eigenvectors corresponding to the controllable poles, i.e., the eigenvalues of A'_{11}.

Proof: Partitioning the eigenvalue decomposition of $A' = S'\Lambda'S'^{-1}$ conformably with A gives

$$\begin{bmatrix} A'_{11} & A'_{12} \\ 0 & A'_{22} \end{bmatrix} = \begin{bmatrix} S'_{11} & S'_{12} \\ 0 & S'_{22} \end{bmatrix} \begin{bmatrix} \Lambda'_1 & 0 \\ 0 & \Lambda'_2 \end{bmatrix} \begin{bmatrix} S'^{-1}_{11} & S'^{-1}_{11} S'_{12} S'^{-1}_{22} \\ 0 & S'^{-1}_{22} \end{bmatrix} \tag{3.20}$$

Here the eigenvectors $\binom{S'_{11}}{0}$ correspond to the eigenvalues Λ'_1 of A'_{11}, and the eigenvectors $\binom{S'_{12}}{S'_{22}}$ correspond to the eigenvalues Λ'_2 of A'_{22}. From the argument leading to theorem 3.5, we know that any x' in the controllable subspace of $(A',\ B')$ must be of the form $x' = \binom{x'_1}{0}$ where x'_1 is an $(m \times 1)$ vector. Since the $(m \times m)$ matrix S'_{11} is nonsingular, there exists an $(m \times 1)$ vector z such that for any x'_1,

$$\begin{bmatrix} x'_1 \\ 0 \end{bmatrix} = \begin{bmatrix} S'_{11} \\ 0 \end{bmatrix} z$$

Therefore $\binom{S'_{11}}{0}$ is a basis for the controllable subspace of (A', B'). ▮

The following lemma is a consequence of the preceding one.

Lemma 3.2: Consider the system $x_{t+1} = Ax_t + Bu_t$, and a controllability canonical form for it, $x'_{t+1} = A'_t + B'u_t$ where $A' = T^{-1}AT$ and $B' = T^{-1}B$, where T is chosen as is described in theorem 3.6. Then the controllable subspace of the pair (A, B) is spanned by the eigenvectors corresponding to the controllable poles of A, which recall are defined as the eigenvalues of A'_{11}.

Proof: We have $A = TA'T^{-1}$, where T is chosen as described in theorem 3.5. We also have the eigenvalue decomposition of $A', A' = S'\Lambda S'^{-1}$. Combining these, we have

$$A = (TS')\Lambda'(TS')^{-1} \tag{3.21}$$

so that (TS') is the matrix of eigenvectors of A. Using the partitioning (3.20) in (3.21) we have

$$A = (T_1 T_2)\begin{bmatrix} S'_{11} & S'_{12} \\ 0 & S'_{22} \end{bmatrix}\begin{bmatrix} \Lambda'_1 & 0 \\ 0 & \Lambda'_2 \end{bmatrix}\left[(T_1 T_2)\begin{bmatrix} S'_{11} & S'_{11} \\ 0 & S'_{22} \end{bmatrix}\right]^{-1}$$

or

$$A = (T_1 S'_{11}, T_1 S'_{12} + T_2 S'_{22})\begin{bmatrix} \Lambda'_1 & 0 \\ 0 & \Lambda'_2 \end{bmatrix}(T_1 S'_{11}, T_1 S'_{12} + T_2 S'_{22})^{-1}$$

Here $T_1 S'_{11}$ are the eigenvectors of A corresponding to the controllable poles Λ'_1, which are the eigenvalues of A'_{11}. The $(n \times m)$ matrix T_1 is a basis for the controllable subspace of $(A,\ B,)$ while the $(m \times m)$ matrix S'_{11} is nonsingular. Therefore, $T_1 S'_{11}$ is a basis for the controllable subspace of $(A,\ B)$. ∎

The following lemma establishes that the property of stabilizability is not disturbed by the application of a nonsingular transformation of the state space.

Lemma 3.3: Consider the system

$$x_{t+1} = Ax_t + Bu_t$$

and let V be *any* nonsingular $(n \times n)$ matrix. Consider the transformed system

$$x'_{t+1} = VAV^{-1}x'_t + VBu_t$$

$$\text{or} \qquad x'_{t+1} = A'x'_t + B'u_t$$

where $x'_t = Vx_t$, where $A' = VAV^{-1}$ and $B' = VB$. Then the pair $(A,\ B)$ is stabilizable if and only if the pair (A', B') is stabilizable.

Proof: Consider the eigenvalue decomposition of $A = S\Lambda S^{-1}$. First partition A as

$$A = (S_1\ \ S_2)\begin{bmatrix} \Lambda_1 & 0 \\ 0 & \Lambda_2 \end{bmatrix}(S_1\ \ S_2)^{-1} \tag{3.22}$$

where the diagonal matrix Λ_1 consists of the *unstable* poles of A and the columns of S_1 are the eigenvectors corresponding to Λ_1. Alternatively, partition the eigenvalue decomposition of A as

$$A = (\tilde{S}_1 \tilde{S}_2) \begin{bmatrix} \tilde{\Lambda}_1 & 0 \\ 0 & \tilde{\Lambda}_2 \end{bmatrix} (\tilde{S}_1 \tilde{S}_2)^{-1} \tag{3.23}$$

where the diagonal matrix $\tilde{\Lambda}_1$ contains the *controllable* poles of A and $\tilde{S}_1$ is an $(n \times m)$ matrix whose columns are the eigenvectors corresponding to $\tilde{\Lambda}_1$. Evidently, the pair $(A,\ B)$ is stabilizable if and only if each of the eigenvalues of Λ_1 also appears in $\tilde{\Lambda}_1$, so that the linear subspace spanned by $\tilde{S}_1$ is included in the subspace spanned by $\tilde{S}_1$. Using $A' = VAV^{-1}$, we have corresponding to (3.22) and (3.23)

$$A' = (VS_1 \;\; VS_2) \begin{bmatrix} \Lambda_1 & 0 \\ 0 & \Lambda_2 \end{bmatrix} (VS_1 \;\; VS_2)^{-1} \tag{3.24}$$

$$A' = (V\tilde{S}_1 \;\; V\tilde{S}_2) \begin{bmatrix} \tilde{\Lambda}_1 & 0 \\ 0 & \tilde{\Lambda}_2 \end{bmatrix} (V\tilde{S}_1 V \;\; \tilde{S}_2)^{-1} . \tag{3.25}$$

Since V is nonsingular, the linear subspace spanned by VS_1 is included in the linear subspace spanned by $V\tilde{S}_1$ if and only if the linear subspace spanned by S_1 is included in the linear subspace spanned by $\tilde{S}_1$. This proves that $(A,\ B)$ is stabilizable if and only if $(A',\ B')$ is stabilizable.

The following theorem provides a useful necessary and sufficient condition for a system (A, B) to be stabilizable.

Theorem 3.12: Consider the linear system

$$x_{t+1} = Ax_t + Bu_t$$

Transform it into the controllability canonical form

$$x_{t+1} = \begin{bmatrix} A'_{11} & A'_{12} \\ 0 & A'_{22} \end{bmatrix} x'_t + \begin{bmatrix} B'_1 \\ 0 \end{bmatrix} u_t$$

where the pair (A'_{11}, B'_{11}) is completely controllable. Then the system $(A,\ B)$ is stabilizable if and only if the matrix A'_{22} is stable.

Proof: From lemma 3.3, it suffices to prove that the pair $(A',\ B')$ is stabilizable if and only if A'_{22} is stable. Partitioning the eigenvalue decomposition of A' we have

$$\begin{bmatrix} A'_{11} & A'_{12} \\ 0 & A'_{22} \end{bmatrix} = \begin{bmatrix} S'_{11} & S'_{12} \\ 0 & S'_{22} \end{bmatrix} \begin{bmatrix} \Lambda'_1 & 0 \\ 0 & \Lambda'_2 \end{bmatrix} \begin{bmatrix} S'^{-1}_{11} & -S'^{-1}_{11} S'_{12} S'^{-1}_{22} \\ 0 & S'^{-1}_{22} \end{bmatrix}$$

A basis for the controllable subspace of $(A',\ B')$ is formed by the eigenvectors $\begin{pmatrix} S'_{11} \\ 0 \end{pmatrix}$ corresponding to the controllable poles Λ'_1. If any eigenvalue $\bar{\lambda}_2$ in Λ'_2 exceeds unity in modulus, then the system cannot be stabilizable, for then the eigenvector corresponding to $\bar{\lambda}_2$ would not belong to the stable subspace. Conversely, if all of the eigenvalues in Λ'_2 and less than unity in modulus, then the unstable subspace is contained in the controllable subspace, a basis for which is $\begin{pmatrix} S'_{11} \\ 0 \end{pmatrix}$.

Therefore, $(A,\ B)$ and its controllability canonical representation (A', B') are stabilizable if and only if A'_{22} is a stable matrix. ▮

5. An Example

As an example to illustrate these concepts, consider the following problem. A firm wants to maximize

$$\sum_{t=0}^{\infty} \beta^t \left[s_t k_t - \frac{f}{2} k_t^2 - J_t (k_{t+1} - k_t)^2 \right] \quad , \quad f > 0 \ , \ 0 < \beta < 1.$$

$0 < \beta < 1, k_0$ given, s_0, s_{-1} given, J_0, J_{-1} given, where k_t is the stock of a factor at time t, J_t is the relative price of capital, and s_t is a shock to technology. We assume that J_t and s_t follow the laws of motion

$$s_t = \rho_1 s_{t-1} + \rho_2 s_{t-2}$$

$$J_t = \mu_1 J_{t-1} + \mu_2 J_{t-2}$$

where we assume that the zeroes of the polynomials

$$1 - \rho_1 z - \rho_2 z^2 = 0 \tag{3.26}$$

and

$$1 - \mu_1 z - \mu_2 z^2 = 0 \tag{3.27}$$

both lie outside the unit circle; i.e. if z_0 solves (3.26) then $\mid z_0 \mid > 1$.

We desire to describe a state-space representation of the linear system the firm is trying to control. We define the state vector x_t and control vector u_t as

$$x_{t+1} = \begin{bmatrix} k_{t+1} \\ s_{t+1} \\ s_t \\ J_{t+1} \\ J_t \end{bmatrix}, \quad u_t = (k_{t+1} - k_t)$$

The system can then be written

$$\begin{bmatrix} k_{t+1} \\ s_{t+1} \\ s_t \\ J_{t+1} \\ J_t \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & \rho_1 & \rho_2 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \mu_1 & \mu_2 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} k_t \\ s_t \\ s_{t-1} \\ J_t \\ J_{t-1} \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} u_t \tag{3.28}$$

We claim that the system (3.28) is in controllability canonical form. First, notice that the matrices corresponding to A and B have zeros in the proper places.

Next, we calculate the controllability matrix

$$\begin{aligned} P_{11} &= [B_1', A_{11}' B_1', \ldots, A_{11}'^{n-1} B_1'] \\ &= [1], \end{aligned}$$

which is evidently of rank 1. Therefore the pair $(A_{11}', B_1') = (1, 1)$ is controllable. So we conclude that the system (3.28) is in controllability canonical form.

We also claim that the system (3.28) is stabilizable. To establish this claim, we must show that the eigenvalues of A_{22}' are less than unity, where

$$A_{22} = \begin{bmatrix} \rho_1 & \rho_2 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & \mu_1 & \mu_2 \\ 0 & 0 & 1 & 0 \end{bmatrix} \tag{3.29}$$

It can be established readily that the four eigenvalues of A_{22}' given by (3.29) equal the *reciprocals* of the four roots of (3.26) and (3.27). For the characteristic equation of A_{22}' is

$$|A_{22}' - \lambda I| = 0$$

which turns out to be

$$(\lambda^2 - \rho_1 \lambda - \rho_2)(\lambda^2 - \mu_1 \lambda - \mu_2) = 0 . \tag{3.30}$$

Setting $z = \lambda^{-1}$ in (3.26) and (3.27) and multiplying (3.26) and (3.27) together gives

$$(1 - \rho_1\lambda^{-1} - \rho_2\lambda^{-2})(1 - \mu_1\lambda^{-1} - \mu_2\lambda^{-2}) = 0$$

or

$$(\lambda^2 - \rho_1\lambda - \rho_2)(\lambda^2 - \mu_1\lambda - \mu_2) = 0$$

The last equation is equivalent with the characteristic equation (3.30) of A'_{22}. Thus, the conditions that the zeros of (3.26) and (3.27) lie outside the unit circle are sufficient to assure that our system is stabilizable.

The reader should convince himself that the system (3.28) is neither controllable nor stable.

Appendix to Chapters 3 & 4

The following definitions and theorems are about the $m \times n$ matrix A and the equation system $Ax = b$ where x is an $n \times i$ vector and b is an $(m \times l)$ vector. Let the rank of A be $k \leq \max(m, n)$.

Theorem B1: There exists a solution to $Ax = b$ if and only if $k = m$. In this case, the columns of A span R^m.

Theorem B2: The $(m \times n)$ matrix A has a right inverse which is an $n \times m$ matrix R satisfying $AR = I_m$, if and only if the rank of $A = m$.

Notice that in the case in which the rank of A is m, one solution of the equation system is $x = Rb$.

Theorem B3: The system $Ax = b$ has at most one solution if and only if $k = n$. In this case, the columns of A are linearly independent.

Theorem B4: The $(m \times n)$ matrix A has a left inverse, which is an $n \times m$ matrix L satisfying $LA = I_n$, if and only if the rank of A equals n.

See Gilbert Strang [p. 71] or Noble and Daniel [pp. 96-97] for proofs of these theorems.

Definition B1: The *null space* of the $(m \times n)$ matrix A, denoted $\mathcal{N}(A)$, is the set of all $(n \times 1)$ vectors x that satisfy $Ax = 0$. The null space is a linear subspace of R^n.

Definition B2: The *range space* or *column space* of A is the set of al $(m \times 1)$ vectors y such that $Ax = y$ for some $x \in R^n$. The range space is a linear subspace of R^n.

Definition B3: *The range space of* A^T or the *row space* of A is the set of all $(m \times 1)$ vectors c that satisfy $A^T z = c$ for some $z \in R^m$. The range space of A^T is a linear subspace of R^n.

Definition B4: Given a subspace of V of R^n, the space of all vectors orthogonal to V is called the *orthogonal complement* of V.

Theorem B5: The null space of A equals the orthogonal complement of the range space of A^T.

Note that the dimension of the range space of A^T is k, while the dimension of the null space of A is $n - k$.

Elementary row operations on a matrix consist of:

(a) Interchange of two rows

(b) Multiplication of any row by a nonzero scalar.

(c) Replacement of the i^{th} row by the sum of the i^{th} row and p times the j^{th} row, (j is not equal to i).

Performing sequence of elementary row operations on A amount to premultiplying it by a non singular matrix. In particular, each elementary row operation on A amounts to premultiplying by the nonsingular matrix that is obtained by performing the same elementary row operation on the identity matrix.

The following theorem is useful to forming a basis for the range space of A.

Theorem B6: Let a series of elementary row operations transform an $(m \times n)$ matrix A into a matrix B. Then a given collection of columns of A is linearly independent (dependent) if and only if the corresponding columns of B are linearly independent (dependent).

Proof: See Noble and Daniel [p. 126].

This theorem is useful in conjunction with the row echelon form in constructing a basis for the column space of A. The row echelon form (see Noble and Daniel) is obtained from A by a series of elementary row operations. In the row echelon form, there are k columns (where $k = rank(A)$) which are the unit vectors $e, \ldots, e_k$. These unit vectors appear in the columns number $c_1, c_2, \ldots, c_k$, with $c_1 < c_2 \leq \ldots \leq c_k$. The last $(m-k)$ rows of the row echelon form are zero, while the first k rows are nonzero. (See Noble and Daniel, p. 88, for more details.)

The preceding theorem implies that a basis for A can be found as follows. Reduce A to row echelon form B. Let the unit vectors in B appear in columns $c_1, \ldots, c_k$. Then the columns number $c_1, c_2, \ldots, c_k$ in A form a basis for the range space of A.

A basis for the null space of A can be constructed as follows. First reduce A to row echelon form by a sequence of elementary row operations, representable by premultiplication of A by the nonsingular matrix E. Note that solutions of $Ax = 0$ are equivalent with solutions of $EAx = 0$. The row echelon form EA has k unit column vectors which can be chosen to be the first k columns of EA by suitably renumbering the variables. Then the row echelon form EA can be written,

$$\begin{bmatrix} I_k & B_{k\times(n-k)} \\ 0_{(m-k)\times k} & 0_{(m-k)\times(n-k)} \end{bmatrix}$$

and $EAx = 0$ can be written:

$$\begin{bmatrix} I_k & B \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

where x_1, is $k \times 1$ and x_2 is $(n-k) \times 1$. Let I_{n-k} be the $(n-k) \times (n-k)$ identity matrix. Then a basis for the null space of A is given by the columns of the $n \times (n-k)$ matrix

$$\begin{bmatrix} -B \\ I_{n-k)} \end{bmatrix}.$$

See Noble and Daniel [pp. 159-160] for more details.

An alternative method of constructing a basis for the null space of A builds upon the fact that the range space of A^T and the null space of A are orthogonal complements. We can construct an orthogonal basis for the range space of A simply by using the Gram-Schmidt orthogonalization procedure. Let $a_1, a_2, \ldots, a_n$ be the n columns of the $(m \times n)$ matrix A. Let the inner product of two vectors y and z in R^m be defined as $(y, z) = \sum_{i=1}^{m} y_i z_i$. Let the norm of y be $||y|| = (y, y)^{\frac{1}{2}}$. Then we recursively form:

$$v_1 = a_1, \quad x_1 = \frac{v_1}{||v_1||}$$

$$v_2 = a_2 - (y_1, a_2) \cdot x_1, \ x_2 = \frac{v_2}{||v_2||}$$

$$\vdots$$

$$v_r = a_r - (x_{r-1}, a_r)X_{r-1} - (x_{r-2}, a_r)\, x_{r-2} - (x_1, a_r) \cdot x_1, \ x_r = \frac{v_r}{||v_r||}$$

If at some step v_r is identically zero, it indicates that a_r is linearly dependent on the preceding $a_1, \ldots, a_{r-1}$. When a null vector v_j is produced by the procedure, omit it and continue until k vectors have been obtained. These k vectors form an orthonormal basis for the range space of A.

To construct a basis for the null space of A, which has dimension $(n-k)$, we first use the Gram-Schmidt orthogonalization procedure to construct an orthonormal basis for the range space of A^T. This is a set of k vectors, $v_1, \ldots, v_k$, which spans a k dimensional subspace of R^n. Next take vectors from the $m \times n$ identity matrix. $I_n = e_1, e_2, \ldots, e_n$, and continue with the Gram-Schmidt orthogonalization procedure until an additional $(n-k)$ orthogonal vector $v_{k+1}, \ldots, v_n$ are found. (This procedure will encounter k vectors that are indentically zero, and are to be omitted.) The vectors $v_{k+1}, \ldots, v_n$ form an orthonormal basis for the null space of A.

For a discussion of the Gram-Schmidt procedure, see Noble and Daniel [pp. 138–139].

Chapter 4

Reconstructibility and Detectability

1. Introduction

This chapter describes the concepts of reconstructibility and detectability. As we shall see repeatedly, what the concept of controllability is to the linear regulator problem, the concept of reconstructibility is to the filtering problem and *vice versa*. Also, what the concept of stabilizability is to the control problem, the concept of detectability is to the filtering problem. Thus, the theorems stated in this chapter will closely resemble those stated in the previous chapter.

2. Reconstructibility

Consider the linear time invariant system

$$x_{t+1} = Ax_t + Bu_t \tag{4.1}$$

$$y_t = Cx_t \tag{4.2}$$

Here x_t is an $(n \times 1)$ vector of *state* variables, u_t is a $(k \times 1)$ vector of inputs or *controls*, and y_t is an $(\ell \times 1)$ vector of *observed* or *output* variables. The matrix A is dimensioned $(n \times n)$, B is $(n \times k)$, and C is $(\ell \times n)$.

Given an initial condition x_{t_0}, the solution of the state difference equation (4.1) is

$$x_t = A^{t-t_0}x_{t_0} + \sum_{s=t_0}^{t-1} A^{t-s-1}Bu_s, \quad t > t_0 \tag{4.3}$$

Using (4.2) in conjunction with (4.3), we obtain the following expression for $y_t, t > t_0$:

$$y_t = CA^{t-t_0}x_{t_0} + \sum_{s=t_0}^{t-1} CA^{t-s-1}Bu_s \tag{4.4}$$

Let $y(t;\ t_0, x_0, \tilde{u}_t)$ denote the response of the output variables of the system over $t \geq t_0$ with initial condition x_{t_0} and control vector $u_t = \tilde{u}_t,\ t \geq t_0$. That is, from (4.4), we define

$$y(t;\ t_0,\ x_0,\ \tilde{u}_t) = CA^{t-t_0}x_{t_0} + \sum_{s=t_0}^{t-1} CA^{t-s-1}B\tilde{u}_s. \tag{4.5}$$

We are now in a position to define the important concept of reconstructibility.

Definition 4.1: The system (4.1)-(4.2) is said to be *reconstructible* or *completely reconstructible* if for all t_1 there exists a t_0 with $-\infty < t_0 < t_1$, such that the condition of identical output variables, namely $y(t;\ t_0,\ x_{t_0},\ \bar{u}_t) = y(t;\ t_0,\ x'_{t_0},\ \bar{u}_t)$, $t_0 \le t \le t_1$, for all input sequences $\bar{u}_t, t_0 \le t \le t_1$, implies that $x_{t_0} = x'_{t_0}$.

Thus, the system is completely reconstructible if, the initial state of the system can be inferred from observations on the controls and the output vectors alone, given a long enough history of observations. If the system is reconstructible, there exists a finite $t_0 < t_1$ such that the initial state x_{t_0} is uniquely determined given knowledge of the output y_t and input u_t sequences. Once x_{t_0} is known, equation (4.3) can then be used to compute x_t for all $t > t_0$.

Theorem 4.1: The system is completely reconstructible if and only if for all t_1 there exists a t_0 with $-\infty < t_0 < t_1$, such that $y(t,\ t_0,\ x_{t_0},\ 0) = 0$ for $t_0 \le t \le t_1$ implies that $x_{t_0} = 0$.

The theorem asserts that the system is reconstructible if and only if, given an input path consisting entirely of zero controls, zero output implies that the initial state is zero.

We now state a theorem that is useful in developing necessary and sufficient conditions for reconstructibility.

Proof: We first show that if the system is reconstructible, there exists a finite $t_0 < t_1$ such that $y(t;\ t_0,\ x_{t_0},\ 0) = 0$ implies that $x_{t_0} = 0$. Suppose that the system is reconstructible. With zero input vector u_t we have from (4.4) that

$$y_t = CA^{(t-t_0)}x_{t_0},\ t_0 \le t \le t_1$$

Since the system is assumed to be reconstructible, and since an initial condition of $x_{t_0} = 0$ gives rise to a zero output y_t, $t_0 \le t \le t_1$, the above equation implies that $x_{t_0} = 0$. This proves half of the theorem: if the system is reconstructible, there exists a finite $t_0 < t_1$ such that $y(t;\ t_0,\ x_{t_0}, 0) = 0$ for $t_0 \le t \le t_1$ implies that $x_{t_0} = 0$.

To prove the other half, assume that there exists a $t_0 \le t_1$ such that $y(t;\ t_0,\ x_{t_0},\ 0) = 0$ for all $t_0 \le t \le t_1$ implies that $x_{t_0} = 0$. By (4.4), we have that $y(t;\ t_0,\ x_{t_0},\ u_t) = y(t;\ t_0,\ x'_{t_0},\ u_t)$ for all $t_0 \le t \le t_1$ is equivalent with $CA^{t-t_0}x_{t_0} = CA^{t-t_0}x'_{t_0}$ for all $t \epsilon [t_0,\ t_1]$. This is equivalent with

$$CA^{(t-t_0)}(x_{t_0} - x'_{t_0}) = 0 \quad \text{for all} \quad t_0 \le t \le t_1 \tag{4.6}$$

But, by hypothesis, there exists a $t_0 < t_1$ such that $y(t;\ t_0,\ x_{t_0}, 0) = 0$ for all $t \varepsilon [t_0,\ t_1]$ implies that $x_{t_0} = 0$. But (4.6) asserts precisely that $y(t,\ t_0,\ x_{t_0} - x'_{t_0},\ 0) = 0$. Therefore $x_{t_0} - x'_{t_0} = 0$ or $x_{t_0} = x'_{t_0}$. Therefore the system is completely reconstructible. ∎

The following theorem states a necessary and sufficient condition for the system to be reconstructible.

Theorem 4.2: The system defined by the pair of matrices (A, C), where A is $n \times n$ and C is $\ell \times n$, is completely reconstructible if and only if the row vectors of the $(n\ell \times n)$ reconstructibility matrix

$$Q = \begin{bmatrix} C \\ CA \\ CA^2 \\ \cdots \\ CA^{n-1} \end{bmatrix}$$

span the n-dimensional space, i.e., if and only if the rank of Q is n.

Proof: By theorem 4.1, the condition that the system be reconstructible is equivalent with the condition that for each t, there exists a finite $t_0 < t_1$ such that $y(t,\ t_0,\ x_{t_0},\ 0) = 0$ implies that $x_{t_0} = 0$. This is equivalent with the existence of a finite $t_0 < t_1$ such that $CA^{(t_1-t_0)}x_{t_0} = 0$ for $t_0 \le t \le t_1$ implies that $x_{t_0} = 0$. Writing this in matrix form gives the requirement that the quality

$$\begin{bmatrix} C \\ CA \\ \vdots \\ CA^{t_1-t_0} \end{bmatrix} x_{t_0} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

for some $t_0 < t_1$ implies $x_{t_0} = 0$. We shall show that it is sufficient to take $t_1 - t_0 = n - 1$ or $t_0 = t_1 - n + 1$.

Thus, take $t_1 - t_0 = n - 1$, and consider the system of linear equations

$$0 = \begin{bmatrix} y_{t_0} \\ y_{t_0+1} \\ \vdots \\ y_{t_0+n-1} \end{bmatrix} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} x_{t_0}$$

or

$$0 = Qx_{t_0}.$$

Suppose that the rank of Q is n. Then it follows from a theorem in linear algebra that $Qx_{t_0} = 0$ implies that $x_{t_0} = 0$. (Recall that the $(\ell n \times n)$ matrix Q has a left inverse satisfying $LQ = I$ if and only if its rank if n. Also, if Q has a left inverse, the solution of $Qx_{t_0} = 0$ is unique, if a solution exists. Since $x_{t_0} = 0$ is a solution, it is the unique solution when Q has rank n.) Therefore, if the rank of Q is n, the system is reconstructible.

To show the converse, first note as an implication of the Cayley-Hamilton theorem that the rows of CA^h are linearly dependent on the rows of Q for $h \geq n$. Therefore, for any $t_0 < t$, the rank of

$$\begin{bmatrix} C \\ CA \\ \vdots \\ CA^{t_1 - t_0} \end{bmatrix}$$

is less than or equal to the rank of Q (equal for $t_1 - t_0 \geq n - 1$). Now if Q has rank less than n, there exists an $x_{t_0} \neq 0$ such that

$$Qx_{t_0} = 0.$$

Further, by the preceding argument, if the rank of Q is less than n, there exists an $x_{t_0} \neq 0$ which solves

$$\begin{bmatrix} C \\ CA \\ \vdots \\ CA^{(t_1 - t_0)} \end{bmatrix} x_{t_0} = 0$$

for any finite $t_0 < t_1$. Therefore, if the rank of Q is less than n, the pair (A, C) is not reconstructible. ∎

3. Examples

(a) Consider the system

$$\theta_{t+1} = \rho\theta_t + \varepsilon_{1t+1}$$

$$y_t = \theta_t + \varepsilon_{2t}$$

where $(\varepsilon_{1t+1}, \varepsilon_{2t})$ is a white noise vector satisfying

$$E \begin{bmatrix} \varepsilon_{1t+1} \\ \varepsilon_{2t} \end{bmatrix} \begin{bmatrix} \varepsilon_{1t+1} \\ \varepsilon_{2t} \end{bmatrix} = \begin{bmatrix} V_1 & V_3 \\ V_3^T & V_2 \end{bmatrix}$$

Here θ_t is a hidden state variable, and y_t is the observed output. For this system, we set $x_t = \theta_t, A = \rho, C = 1$. So long as $\rho \neq 1$, $(A,\ C)$ is reconstructible

(b) Consider the system

$$\theta_{t+1} = \rho_1\theta_t + \rho_2\theta_{t-1} + \varepsilon_{1t+1}$$

$$y_t = \theta_t + \varepsilon_{2t}$$

where $(\varepsilon_{t+1}, \varepsilon_t)$ is again a vector white noise. The system can be written

$$x_{t+1} = \begin{bmatrix}\theta_{t+1}\\ \theta_t\end{bmatrix} \begin{bmatrix}\rho_1 & \rho_2\\ 1 & 0\end{bmatrix}\begin{bmatrix}\theta_t\\ \theta_{t-1}\end{bmatrix} + \begin{bmatrix}\varepsilon_{1t+1}\\ 0\end{bmatrix}$$

$$y_t = (1\ 0)\,x_t + \varepsilon_{2t}.$$

For this system, we take $A = \begin{bmatrix}\rho_1 & \rho_2\\ 1 & 0\end{bmatrix}$, $C = (1\ 0)$. The observability matrix Q is $Q = \begin{bmatrix}1 & 0\\ \rho_1 & \rho_2\end{bmatrix}$. So long as $\rho_2 \neq 0$, the pair $(A,\ C)$ is reconstructible.

(c) Consider the system

$$\theta_{t+1} = \rho\theta_t + \varepsilon_{1t+1}$$

$$\phi_{t+1} = \alpha\phi_{t+1} + \varepsilon_{2t+1}$$

$$y_t = c_1\theta_t + c_2\phi_t + \varepsilon_{3t}$$

where $(\varepsilon_{1t+1},\ \varepsilon_{2t+1},\ \varepsilon_{3t})$ is a vector white noise. For this system, define

$$x_t = \begin{bmatrix}\theta_t\\ \phi_t\end{bmatrix},\ A = \begin{bmatrix}\rho & 0\\ 0 & \alpha\end{bmatrix},\ C = (c_1\ c_2).$$

We then have the observability matrix

$$Q = \begin{bmatrix}c_1 & c_2\\ c_1\rho & c_2\alpha\end{bmatrix}$$

The system is observable unless $c_1\ c_2 = 0$ or $\rho = \alpha$.

(d) Consider the maximum problem:

$$maximize \quad \sum_{t=t_0}^{t_1-1} -x_t^T C^T C x_t + u_t^T Q_t$$

subject to $x_{t+1} = Ax_t + Bu_t$. As we shall see in the next chapter, it is of interest to determine the reconstructibiltiy status of the pair (A, C). We invite the reader to check the reconstructibility status of the (A, C) pairs for the following problems:

(i) The "transformed" consumption problem given above on page 2-8, 2-9.

(ii) The capital accumulation problem given in pages 3: 25-28.

4. The Reconstructibility Canonical Form

We next define the concept of the unreconstructible subspace.

Definition 4.2: The *unreconstructible subspace* of the system (4.1)-(4.2) is the linear subspace of states x_{t_0} for which $y(t;\ x_{t_0},\ t_0,\ 0) = 0, t \geq t_0$.

Theorem 4.3: The unreconstructible subspace of the n-dimensional system (4.1)-(4.2) is the null space of the reconstructibility matrix

$$Q = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix}$$

Proof: This follows directly from the machinery in the proof of the previous theorem, and from the definition of the null space of $Q, \mathcal{N}(Q)$, as the set of vectors x such that $Qx = 0$. Thus any initial state vector x_{t_0} in the null space of Q produces an identically zero output in response to a zero input, while any initial state vector x_{t_0} not in the null space of Q produces a non-zero response. ∎

The following lemma describes a characteristic of the unreconstructible subspace of (A, C) which we shall use.

Lemma 4.1: The unreconstructible subspace of (A, C) is invariant under A.

Proof: We must show that if x_{t_0} belongs to $\mathcal{N}(Q)$, then Ax_{t_0} also belongs to $\mathcal{N}(Q)$. By the Cayley-Hamilton theorem, we know that there exist scalars α_k^h such that

$$A^h = \sum_{k=0}^{n-1} \alpha_k^h A^k, h \geq n.$$

It follows that

$$CA^h = \sum_{k=0}^{n-1} \alpha_k^h CA^k, h \geq n.$$

Now let x_{t_0} be in the null space of Q, or

$$\begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} x_{t_0} = 0.$$

Since $CA^n x_{t_0} = \sum_{k=0}^{n-1} \alpha_k^n (CA^k) x_{t_0} = 0$, it follows that

$$QAx_{t_0} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} Ax_{t_0} = \begin{bmatrix} CA \\ CA^2 \\ \vdots \\ CA^n \end{bmatrix} x_{t_0} = 0.$$

Therefore Ax_{t_0} is in the null space of Q. ▮

The next theorem indicates that the state of the system can be determined only to within the addition of an arbitrary vector in the unreconstructible subspace.

Theorem 4.4: Consider the system (4.1)–(4.2). Suppose y_t and u_t are known over an interval $t_0 \leq t \leq t_1$, with $t_1 - t_0 \geq n - 1$.

(a) The initial state of the system at time t_0 is determined to within the addition of an arbitrary vector in $\mathcal{N}(Q)$.

(b) The terminal state at time t_1 is determined to within the addition of a vector of the form $A^{(t_1 - t_0)} \tilde{x}_{t_0}$, where $\tilde{x}_{t_0}$ is an arbitrary vector in $\mathcal{N}(Q)$.

Proof of (a): We must show that if x_{t_0} and x'_{t_0} produce the same output $y_t, t_0 \leq t \leq t_1$, for any input $u_t, t_0 \leq t \leq t_1$, then $x_{t_0} - x'_{t_0}$ belongs to $\mathcal{N}(Q)$. Now, $y(t, t_0, x_{t_0}, u) = y(t; t_0, x'_{t_0}, u)$ for $t_0 \leq t \leq t_1$ is equivalent with $CA^{(t-t_0)} x_{t_0} = CA^{(t-t_0)} x'_{t_0}$ for $t_0 \leq t \leq t_1$. This is equivalent with $CA^{t-t_0}(x_{t_0} - x'_{t_0})$ for all $t_0 \leq t \leq t_1$. This implies that $x_{t_0} - x'_{t_0}$ belongs to $\mathcal{N}(Q)$.

Proof of (b): The addition of an arbitrary vector x''_{t_0} in $\mathcal{N}(Q)$ to the initial state results in the addition to the output of $CA^{(t-t_0)} x''_{t_0}$. Since x''_{t_0} belongs to $\mathcal{N}(Q)$, the addition is zero for all $t \leq t_0$. ▮

The next theorem represents the structure of the system in a way that we shall find very useful.

Theorem 4.5: Consider the n^{th} - order time invariant system

$$x_{t+1} = Ax_t + Bu_t$$

$$y_t = Cx_t$$

Form a nonsingular transformation matrix $U = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix}$ where the m rows of U_1 form a basis for the m-dimensional ($m \leq n$) subspace spanned by the rows of Q. The $(n - m)$ rows of

U_2 are chosen so that, together with the m rows of U_1, they form a basis for R^n. Define a transformed state vector x'_t by

$$x'_t = Ux_t$$

Then in terms of x'_t the system is represented in *reconstructibility canonical form*

$$x'_{t+1} = \begin{bmatrix} A'_{11} & 0 \\ A'_{21} & A'_{22} \end{bmatrix} x'_t + \begin{bmatrix} B'_1 \\ B'_2 \end{bmatrix} u_t \tag{4.7}$$

$$y_t = (C'_1 \ 0)x'_t \tag{4.8}$$

Here A'_{11} is an $(m \times m)$ matrix, C'_1 is $(\ell \times m), A'_{22}$ is $(n-m) \times (n-m)$, and B_1 is $(m \times k)$. The pair (A'_{11}, C'_1) is completely reconstructible.

Proof: Suppose that *rank* $(Q) = m \leq n$, so that Q possesses m linearly independent rows. This implies that the null space of Q has dimension $(n-m)$. Let the row vectors $f_1, \dots, f_m$ be a basis for the m-dimensional linear space spanned by the row vectors of Q, i.e., the range space of Q^T. Let $f_{m+1}, \dots, f_n$ be $(n-m)$ linearly independent row vectors that together with $f_1, \dots, f_m$ span R^n. Now form

$$U = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix},$$

$$U_1 = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_m \end{bmatrix}, \quad U_2 = \begin{bmatrix} f_{m+1} \\ \vdots \\ f_n \end{bmatrix}$$

Introduce the transformed state vector $x'_t = Ux_t$ so that $x_t = U^{-1}x'_t$. Then the system (4.1)–(4.2) can be written

$$U^{-1}x'_{t+1} = AU^{-1}x'_t + Bu_t$$

$$y_t = CU^{-1}x'_t$$

or

$$x'_{t+1} = UAU^{-1}x'_t + UBu_t \tag{4.9}$$

$$y_t = CU^{-1}x'_t \tag{4.10}$$

Partition U^{-1} conformably with the partition of U, so that

$$U^{-1} = (T_1 \ T_2)$$

where T_1 has m columns and T_2 has $(n-m)$ columns. We have

$$UU^{-1} = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} (T_1 T_2) = \begin{bmatrix} U_1 T_1 & U_1 T_2 \\ U_2 T_1 & U_2 T_2 \end{bmatrix} = \begin{bmatrix} I_m & 0 \\ 0 & I_{n-m} \end{bmatrix}$$

which implies that $U_1 T_2 = 0$. Now the rows of U_1 form a basis for the range space of Q^T, so that any vector x that satisfies $U_1 x = 0$ also satisfies $Qx = 0$. ($U_1 x = 0$ states that x is orthogonal to the range space of Q^T and therefore is in the null space of Q. For recall from linear algebra that the null space of Q is the orthogonal complement of the range space of Q^T.) Since $U_1 T_2 = 0$, it follows that all columns of T_2 are in $\mathcal{N}(Q)$. Because T_2 has $(n-m)$ linearly independent column vectors, and the unreconstructible subspace has dimension $(n-m)$, the column vectors of T_2 form a basis for $\mathcal{N}(Q)$. Therefore, $U_1 x = 0$ for any x belonging to $\mathcal{N}(Q)$.

We can write

$$UAU^{-1} = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} A[T_1 T_2] = \begin{bmatrix} U_1 A T_1 & U_1 A T_2 \\ U_2 A T_1 & U_2 A T_2 \end{bmatrix}$$

$$CU^{-1} = [CT_1, CT_2]$$

All of the column vectors of T_2 are in $N(Q)$. Because $N(Q)$ is invariant under A (by lemma 4.1), the columns of AT_2 are in $\mathcal{N}(Q)$. Therefore, $U_1 A T_2 = 0$. Since the rows of C are rows of Q, and since the columns of T_2 are in $N(Q)$, we also have $CT_2 = 0$. Thus, the above equations become

$$A' = UAU^{-1} = \begin{bmatrix} U_1 A T_1 & 0 \\ U_2 A T_1 & U_2 A T_2 \end{bmatrix} = \begin{bmatrix} A'_{11} & 0 \\ A'_{21} & A'_{22} \end{bmatrix}$$

$$C' = CU^{-1} = (CT_1, 0) = (C'_1, 0).$$

Thus, the system can be written in the form of (4.7) and (4.8),

(4.7) $$x'_{t+1} = \begin{bmatrix} A'_{11} & 0 \\ A'_{21} & A'_{22} \end{bmatrix} x'_t + \begin{bmatrix} B'_1 \\ B'_2 \end{bmatrix} u_t$$

(4.8) $$y(t) = (C'_1 \ 0) x'_t$$

It remains to verify that the pair (A'_{11}, C'_1) is reconstructible. Recall that A'_{11} is $(m \times m)$ and C'_1 is $(\ell \times m)$. First notice that the reconstructibility matrix Q' for the transformed system (4.7)-(4.8) is given by

$$Q' = \begin{bmatrix} C' \\ C'A' \\ \vdots \\ C'A'^{n-1} \end{bmatrix} = \begin{bmatrix} CU^{-1} \\ CAU^{-1} \\ \vdots \\ CA^{n-1}U^{-1} \end{bmatrix} = QU^{-1}$$

Since U^{-1} is nonsingular, it follows that the rank of Q' equals the rank of Q, which is $m \leq n$ by assumption. From the equations $C' = (C'_1,\ 0)$, $A' = \begin{bmatrix} A'_{11} & 0 \\ A'_{21} & A'_{22} \end{bmatrix}$, Q' is calculated to be

$$Q' = \begin{bmatrix} C'_1 & 0 \\ C'_1 A_{11} & 0 \\ \vdots & \vdots \\ C'_1 A_{11}^{n-1} & 0 \end{bmatrix}$$

Since the rank of Q' is m, it follows from the above equation and the Cayley-Hamilton theorem that the $(\ell m \times m)$ matrix

$$Q'_1 = \begin{bmatrix} C'_1 \\ C'_1 A_{11} \\ \vdots \\ C'_1 A_{11}^{m-1} \end{bmatrix}$$

has rank m. Since Q'_1 is the reconstructibility matrix for the pair $(A'_{11},\ C'_1)$, it follows that the pair is completely reconstructible. ∎

The characteristic values of A'_{11} and of A'_{22} are independent of the particular choice of U_1 and U_2. The proof of this assertion uses the same logic that was earlier used to prove theorem 4.5 and will be omitted. The characteristic values of A'_{11} are called the *reconstructible poles*, while those of A'_{22} are called the *unreconstructible poles*. The unreconstructible subspace is spanned by the characteristic vectors corresponding to the unreconstructible poles, while the reconstructible subspace is spanned by the characteristic vectors corrresponding to the reconstructible poles. The proof of this assertion uses the same logic that was used to prove the lemma 3.2 about controllable and uncontrollable subspaces.

5. Detectability

Next we have an important definition.

Definition 4.3 The linear time invariant system (4.1)-(4.2) is said to be *detectable* if its unreconstructible subspace is contained in its stable subspace.

We have the following theorems.

Theorem 4.6: Any stable system is detectable.

Proof: The unstable subspace is empty. ▮

Theorem 4.7: Any completely reconstructible system is detectable.

Proof: The unreconstructible subspace is empty. ▮

Theorem 4.8: Consider the linear time invariant system $x_{t+1} = Ax_t, y_t = Cx_t$. Transform it into the reconstructibility canonical form

$$x'_{t+1} = \begin{bmatrix} A'_{11} & 0 \\ A'_{21} & A'_{22} \end{bmatrix} x'_t$$
$$y_t = (C'_1,\ 0)x'_t$$

where $(A'_{11},\ C'_1)$ is completely reconstructible. Then the system is detectable if and only if A'_{22} is a stable matrix.

Proof: Detectability requires that if x_{t_0} belongs to $\mathcal{N}(Q)$, then $x_t = A^{(t-t_0)}x_{t_0} \to 0$ as $t \to \infty$. If x_{t_0} is in $\mathcal{N}(Q)$, then it has the representation $x_{t_0} = U^{-1}\begin{bmatrix} 0 \\ x'_{2t_0} \end{bmatrix}$, where U^{-1} is the nonsingular transformation matrix defined in theorem 4.5, x'_{2t_0} is an $(n-m)\times 1$ vector, and $U^{-1} = (T_1,\ T_2)$ where the columns of T_1 form a basis for the reconstructible subspace and the columns of T_2 form a basis for the unreconstructible subspace of $(A,\ C)$. We have that

$$x_t = U^{-1}\begin{bmatrix} 0 \\ A_{22}'^{(t-t_0)}x'_{2t_0} \end{bmatrix}$$

Since U^{-1} is nonsingular, $x_t \to 0$ as $t \to \infty$ for all x'_{2t_0} if and only if A'_{22} is a stable matrix. ▮

An alternative proof of this theorem in terms of the eigenvectors of the reconstructible and stable subspaces, can be constructed paralleling the argument in the proof of theorem 3.12 on stabilizability.

We now describe the concept of duality, which is a very useful tool for clarifying the relationship between control and filtering problems.

Definition 4.4: Consider the system

$$(4.11)\qquad \begin{aligned} x_{t+1} &= A x_t + B u_t \\ &(n\times 1)(n\times n)\ (k\times 1) \\ y_t &= C\ x_t \\ &(1\times 1)(1\times n)\ (n\times 1) \end{aligned}$$

The *dual system* is defined as

$$(4.12)\qquad \begin{aligned} x^*_{t+1} &= A^T x^*_t + C^T u^*_t \\ &(n\times 1)(n\times n)(n\times 1)(1\times 1) \\ y^*_t &= B^T x^*_t \\ &(k\times 1)(1\times n) \end{aligned}$$

The following theorem is immediate:

Theorem 4.9: The dual of the dual is the original system.

The following theorem is also immediate:

Theorem 4.10: Consider system (4.11) and its dual (4.12). The following statements are true:

(a) The system (4.11) is completely controllable if and only if the dual (4.12) is completely reconstructible.

(b) The system (4.11) is completely reconstructible if and only if the dual (4.12) is controllable.

(c) System (4.11) is stabilizable if and only if the dual (4.12) is detectable.

(d) System (4.11) is detectable if and only if the dual (4.12) is stabilizable.

Proof: (a) and (b): form the appropriate controllability matrix P and reconstructibility matrix Q.

(c) Transform (4.11) via $x'_t = T^{-1}x_t$ into a controllability canonical form

$$x'_{t+1} = \begin{bmatrix} A'_{11} & A'_{12} \\ 0 & A'_{22} \end{bmatrix} x'_t + \begin{bmatrix} B'_1 \\ 0 \end{bmatrix} u_t$$

$$y_t = (C'_1\ C'_2)x'_t.$$

where $(C_1' \; C_2') = CT$. If the system (4.11) is stabilizable, then (A_{11}', B_1') is controllable and A_{22}' is stable. The dual of the transformed system is

$$\begin{aligned} x_{t+1}'^{*} &= \begin{bmatrix} A_{11}'^T & 0 \\ A_{12}'^T & A_{22}'^T \end{bmatrix} x_t'^{*} + \begin{bmatrix} C_1'^T \\ C_2'^T \end{bmatrix} u_t^{*} \\ y_t^{*} &= (B_1'^T, 0) x_t'^{*} \end{aligned} \tag{4.13}$$

Since A_{11}', B_1' is completely controllable, $A_{11}'^T, B_1'^T$ is completely reconstructible, as can be verified by checking the ranks of the pertinent controllability and reconstructibility matrices. Since A_{22}' is stable, so is $A_{22}'^T$. Therefore, system (4.13) is detectable. By the nonsingular transformation $T^T x_t^{*} = x_t'^{*}$ it can be verified that the system (4.13) is transformed into the dual of the system (4.11). Therefore, since the system (4.13) is also detectable, the dual of the system (4.11) is also detectable. The inverse is easily proved, as is (d). ∎

Chapter 5

Convergence and Stability Theorems For The Optimal Linear Regulator Problem

1. Introduction

This chapter collects the dividends we have earned by investing the last two chapters in the ideas of controllability, stabilizability, reconstructibility, and detectability. We shall use these ideas repeatedly in order to establish convergence and stability theorem for the optimal linear regulator. Actually, we shall earn a double return on our investments, because by repeated appeals to duality, the theorems established in this chapter will be used in chapter 6 to state theorems that apply to the Kalman filter.

2. The Optimal Linear Regulator Problem Again

The following lemma will prove useful:

Lemma 5.1: Consider the quantity

$$J(x_{t_0}) = E_{t_0}\left\{\sum_{t=t_0}^{t_1-1} x_t^T H_t x_t + x_{t_1}^T G_{t_1} x_{t_1}\right\}, \ t_0 < t_1,$$

where the H_t's are given $n \times n$ matrices, where G_{t_1} is a given $n \times n$ matrix, and where $\{x_t\}$ obeys the vector stochastic difference equation

$$x_{t+1} = A_t x_t + \xi_{t+1}.$$

Assume that ξ_{t+1} is an $n \times 1$ vector white noise satisfying

$$E\xi_t = 0 \quad \forall t$$

$$E\xi_t\xi_s^T = 0 \quad \text{for } t \neq s$$

$$E\xi_t\xi_t^T = V_t \quad \forall t,$$

V_t a positive definite matrix. Assume that ξ_t is orthogonal to past x_t, so that

$$Ex_t\xi_s^T = 0 \quad \text{for } t < s.$$

Then

$$J(x_{t_0}) = x_{t_0}^T G_{t_0} x_{t_0} + d_{t_0}$$

where G_{t_0} and d_{t_0} are solutions for t_0 of the difference equations

$$G_{t-1} = A_{t-1}^T G_t A_{t-1} + H_{t-1} \tag{5.1}$$

$$d_{t-1} = d_t + trV_tG_t \tag{5.2}$$

with terminal conditions G_{t_1} and $d_{t_1} = 0$ given.

Proof: Writing out $J(x_{t_0})$ we have

$$\begin{aligned} J(x_{t_0}) = E_{t_0}\{x_{t_0}^T H_{t_0} x_{t_0} + x_{t_0+1}^T H_{t_0+1} x_{t_0+1} + \cdots \\ + x_{t_1-1}^T H_{t_1-1} x_{t_1-1} + x_{t_1}^T G_{t_1} x_{t_1}\}. \end{aligned} \tag{5.3}$$

Using the law of iterated expectations repeatedly, this can be written

$$\begin{aligned} J(x_{t_0}) = E_{t_0}\{x_{t_0}^T H_{t_0} x_{t_0} + E_{t_0+1}\{x_{t_0+1}^T H_{t_0+1} x_{t_0+1} \\ + E_{t_0+2}\{x_{t_0+2}^T H_{t_0+2} x_{t_0+2} + \cdots \\ + E_{t_1-1}^T\{x_{t_1-1}^T H_{t_1-1} x_{t_1-1} + E_{t_1} x_{t_1}^T G_{t_1} x_{t_1}\}\}\}\}. \end{aligned} \tag{5.4}$$

We shall "work backwards", starting by evaluating the terms conditioned on the most information. Since x_t is assumed to be included in the conditioning set at time t, we have

$$E_{t_1} x_{t_1}^T G_{t_1} x_{t_1} = x_{t_1}^T G_{t_1} x_{t_1}.$$

Next, we have

$$\begin{aligned} E_{t_1-1}\{x_{t_1-1}^T H_{t_1-1} x_{t_1-1} + x_{t_1}^T G_{t_1} x_{t_1}\} &= x_{t_1-1}^T H_{t_1-1} x_{t_1-1} \\ &+ E_{t_1-1}\{(A_{t_1-1} x_{t_1-1} + \xi_{t_1})^T G_{t_1} (A_{t_1-1} x_{t_1-1} + \xi_{t_1})\} \\ &= x_{t_1-1}^T H_{t_1-1} x_{t_1-1} + x_{t_1-1}^T A_{t_1-1}^T G_{t_1} A_{t_1-1} x_{t_1-1} + E_{t_1-1}(\xi_{t_1}^T G_{t_1} \xi_{t_1}) \\ &= x_{t_1-1}^T (H_{t_1-1} + A_{t_1-1}^T G_{t_1} A_{t_1-1}) x_{t_1-1} + trV_{t_1} G_{t_1} \\ &= x_{t_1-1}^T G_{t_1-1} x_{t_1-1} + d_{t_1-1}. \end{aligned}$$

Continuing to work backwards leads to the result that for

$$t_0 \le t \le t_1 - 1$$

$$\begin{aligned} E_t\{x_t^T H_t x_t + E_{t+1}\{x_{t+1}^T H_t x_{t+1} + \cdots + E_{t_1} x_{t_1}^T G_{t_1} x_{t_1}\} \cdots\}\} \\ = x_t^T G_t x_t + d_t \end{aligned}$$

where G_t and d_t are the solutions from (1) and (2). ∎

Let R be a $(n \times n)$ negative semidefinite matrix, Q a $(k \times k)$ negative definite matrix and P_{t_1} a given $(n \times n)$ negative semidefinite matrix. Consider the criterion

$$J(x_{t_0}) = E_{t_0}\Big[\sum_{t=t_0}^{t_1-1}(x_t^T R x_t + v_t^T Q v_t) + x_{t_1}^T P_{t_1} x_{t_1}\Big] \tag{5.5}$$

where the system obeys the stochastic difference equation

$$x_{t+1} = Ax_t + Bv_t + \xi_{t+1} \tag{5.6}$$

where ξ_{t+1} is a vector white noise with $E\xi_t\xi_t^T = V_t$. Suppose that v_t is set according to the control law

$$v_t = -F_t x_t \tag{5.7}$$

where $\{F_t\}$ is an arbitrary sequence of $k \times n$ matrices. Substituting (5.7) into (5.6) gives the "closed loop" system equation

$$x_{t+1} = (A - BF_t)x_t + \xi_{t+1}. \tag{5.8}$$

Substituting (5.7) into the criterion function (5.5) gives the following expression:

$$J(x_{t_0}) = E_{t_0}\Big[\sum_{t=t}^{t_1-1} x_t^T\{R + F_t^T Q F_t\}x_t + x_{t_1}^T P_{t_1} x_{t_1}\Big]. \tag{5.9}$$

We can now state the following useful theorem:

Theorem 5.1: Consider the criterion function (5.5) and the system (5.6) operating under the prescribed feedback law (5.7). The criterion function takes the value

$$J(x_{t_0}) = x_{t_0}^T P_{t_0} x_{t_0} + d_{t_0}$$

where P_{t_0} and d_{t_0} are the solutions of

$$P_{t-1} = (A - BF_{t-1})^T P_t (A - BF_{t-1}) + R + F_{t-1}^T Q F_{t-1}$$

$$d_{t-1} = d_t + trV_t P_t$$

with terminal conditions P_{t_1} and $d_{t_1} = 0$ given.

Proof: In lemma 5.1, set $A_{t-1} = (A - BF_{t-1}), G_{t_1} = P_{t_1}, H_t = R + F_t^T Q F_t$. ∎

We can now define an important problem.

Definition: Let $R \leq 0, Q \leq 0, P_{t_1} \leq 0$ be given. Consider the criterion

$$E_{t_0}\Big[\sum_{t=t_0}^{t_1-1} \{x_t^T R x_t + v_t^T Q v_t\} + x_{t_1}^T P_{t_1} x_{t_1}\Big], \tag{5.10}$$

where the system is governed by

$$x_{t+1} = Ax_t + Bv_t + \xi_{t+1}, \tag{5.11}$$

x_{t_0} given, ξ_{t+1} being a vector white noise with $E\xi_t\xi_t^T = V_t$. The problem of maximizing (5.10) subject to (5.11) with respect to choice of

$$F_{t_0}, F_{t_0+1}, \cdots, F_{t_1-1}$$

where $v_t = -F_t x_t$ is called the *optimal linear regulator problem*.

From theorem 5.1, we know that for a *given* sequence $F_{t_0}, F_{t_0+1}, \cdots, F_{t_1-1}$, the criterion (5.10) is given by

$$x_{t_0}^T P_{t_0} x_{t_0} + d_{t_0} \tag{5.12}$$

where P_{t_0} is the solution to

$$P_{t-1} = (A - BF_{t-1})^T P_t (A - BF_{t-1}) + R + F_{t-1}^T Q F_{t-1} \tag{5.13}$$

with P_{t_1} given and

$$d_{t-1} = d_t + tr V_t P_t \tag{5.14}$$

with $d_{t_1} = 0$ given. Our object is to find a sequence of control laws $\{F_{t_0}^o, F_{t_0+1}^o, \cdots, F_{t_1-1}^o\}$ that maximizes $x_{t_0}^T P_{t_0} x_{t_0} + d_{t_0}$ for all x_{t_0}. We make two observations that simplify the task of maximization. First, from (5.14) it follows from the specifications $V_t > 0$ and $P_{t_1} \leq 0$ that d_{t-1} is a monotonically increasing function of d_t and P_t. It immediately follows that d_{t-1} is a monotonically increasing function of $P_t, P_{t+1}, \cdots, P_{t_1-1}$. Second, from (5.13) it follows that given F_{t-1}, P_{t-1} is a monotonically increasing function of P_t. From these observations, it follows that in order to maximize $x_{t_0}^T P_{t_0} x_{t_0} + d_{t_0}$ over $\{F_{t_0}, F_{t_0+1}, \cdots, F_{t_1-1}\}$, the F_t's should be chosen to maximize the matrices $P_t, t = t_0, \cdots, t_1 - 1$, subject to (12) with P_{t_1} given.[1] We say that F_{t-1}^o is the choice of F_{t-1} that maximizes P_{t-1} with P_t^o given, yielding maximized value P_{t-1}^o, if for all other choices of $F_{t-1}, (P_{t-1}^o - P_{t-1}) \geq 0$. The statement $(P_{t-1}^o - P_{t-1}) \geq 0$ means that $P_{t-1}^o - P_{t-1}$ is positive semidefinite, which is equivalent with the statement that $x_{t-1}^T P_{t-1}^o x_{t-1} - x_{t-1}^T P_{t-1} x_{t-1} \geq 0$ for all vectors x_{t-1}.

The preceding observations imply that to maximize $x_{t_0}^T P_{t_0} x_{t_0} + d_{t_0}$ (uniformly in x_{t_0}), it is sufficient to proceed sequentially, working backwards to produce a sequence of maximal $\{P_{t_1-1}^o, P_{t_1-2}, \cdots, P_{t_0}^0\}$. That is, given P_{t_1}, choose F_{t_1-1} to maximize

$$P_{t_1-1} = (A - BF_{t_1-1})^T P_{t_1} (A - BF_{t_1-1}) + R + F_{t_1-1}^T Q F_{t_1-1}. \tag{5.15}$$

Then substitute the optimizing $F_{t_1-1} = F_{t_1-1}^o$ into (5.12) to calculate the maximized values for $P_{t_1-1} = P_{t_1-1}^o$. Next choose F_{t_1-2} to maximize

$$P_{t_1-2} = (A - BF_{t_1-2})^T P_{t_1-1} (A - BF_{t_1-2}) + R + F_{t_1-2}^T Q F_{t_1-2},$$

and so on. So at time $t-1$ we have to choose F_{t-1} to maximize

$$P_{t-1} = (A - BF_{t-1})^T P_t^o (A - BF_{t-1}) + R + F_{t-1}^T Q F_{t-1}$$

or

$$\begin{aligned} P_{t-1} = {} & A^T P_t^o A - A^T P_t^o B F_{t-1} - F_{t-1}^T B^T P_t^o A \\ & + F_{t-1}^T B^T P_t^o B F_{t-1} + R + F_{t-1}^T Q F_{t-1} \end{aligned}$$

or

$$\begin{aligned} P_{t-1} = {} & F_{t-1}^T (B^T P_t^o B + Q) F_{t-1} - F_{t-1}^T B^T P_t^o A \\ & - A^T P_t^o F_{t-1} + R + A^T P_t^o A. \end{aligned}$$

Complete the square by adding and subtracting $(A^T P_t^o B)(B^T P_t^o B + Q)^{-1} (B^T P_t^o A)$ from the right side of the above equation to get

$$\begin{aligned} P_{t-1} = {} & \left[F_{t-1}^T - (A^T P_t^o B)(B^T P_t^o B + Q)^{-1}\right] \left[B^T P_t^o B + Q\right] \\ & \cdot \left[F_{t-1}^T - (A^T P_t^o B)(B^T P_t^o B + Q)^{-1}\right]^T \\ & - (A^T P_t^o B)(B^T P_t^o B + Q)^{-1} (B^T P_t^o A) + R + A^T P_t^o A. \end{aligned} \tag{5.16}$$

[1] Note that since P_t for $t = t_0, \cdots, t_1 - 1$ obeying (5.12) are negative semidefinite, maximizing the P_t's results in minimizing the absolute value of the term d_{t_0}, which is necessarily nonpositive.

Recall that $\left[B^T P_t^o B + Q\right]$ is a negative definite matrix, and that F_{t-1} appears only in the first quadratic form on the right side of the equation. Therefore, to maximize P_{t-1} given P_t^o, F_{t-1} should be chosen so that

$$F_{t-1}^{oT} = (A^T P_t^o B)(B^T P_t^o B + Q)^{-1}$$

or

$$F_{t-1}^o = (B^T P_t^o B + Q)^{-1} B^T P_t^o A. \tag{5.17}$$

Notice that this choice of F_{t-1} makes the first negative definite quadratic form on the right side of (5.16) vanish. Substituting the optimal F_{t-1}^o from (5.17) into (5.16) then gives the following equation for P_{t-1}^o:

$$P_{t-1}^o = A^T P_t^o A + R - A^T P_t^o B(B^T P_t^o B + Q)^{-1} B^T P_t^o A. \tag{5.18}$$

This equation is known as the *matrix Riccati difference equation*.

We summarize these results in a theorem.

Theorem 5.2: Consider the optimal linear regulator problem, to maximize

$$E_{t_0}\left[\sum_{t=t_0}^{t_1-1}\{x_t^T R x_t + v_t^T Q v_t\} + x_{t_1}^T P_{t_1} x_{t_1}\right] \tag{5.10}$$

subject to x_{t_0} given, $R \leq 0, Q < 0, P_{t_0} \leq 0$; where the system dynamics are given by

$$x_{t+1} = Ax_t + Bv_t + \xi_{t+1}\ , \tag{5.11}$$

where ξ_{t+1} is a vector white noise with $E\xi_t\xi_t^T = V_t$. The maximization of (5.10) is carried out over the parameters of feedback rules F_t in

$$v_t = -F_t x_t, \quad t = t_0, t_0+1, \cdots, t_1 - 1.$$

For an *arbitrary* $\{F_t\}$ sequence, the value of the criterion (5.10) is $x_{t_0}^T P_{t_0} x_{t_0} + d_{t_0}$ where P_{t_0} and d_{t_0} are the solutions to the difference equation

$$P_{t-1} = (A - BF_{t-1})P_t(A - BF_{t-1}) + R + F_{t-1}^T Q F_{t-1}$$

$$d_{t-1} = d_t + trV_t P_t$$

with terminal conditions P_{t_1} and $d_{t_1} = 0$ given. The *optimal* choice of F_t's is given by

$$F_t^o = (B^T P_{t+1}^o B + Q)^{-1} B^T P_{t+1}^o A \quad t = t_0, t_0+1, \cdots, t_1 - 1 \tag{5.17}$$

where P_t^o is the solution of the matrix Riccati difference equation

$$P_{t-1}^o = A^T P_t^o A + R - A^T P_t^o B (B^T P_t^o B + Q)^{-1} B^T P_t^o A \tag{5.18}$$

with terminal condition P_{t_1} given. The matrices $\{P_t^o\}$ are negative semidefinite. When the optimal feedback rules are used, the criterion function attains the value

$$x_{t_0}^T P_{t_0}^o x_{t_0} + d_{t_0}.$$

The matrix $P_{t_0}^o$ maximizes P_{t_0} with respect to $\{F_t, t = t_0, t_0+1, \cdots, t_1 - 1\}$ over the class of all matrices $P_{t_0}^o$ that satisfy (5.16) with terminal condition P_{t_1} given. This concludes the statement of the theorem.

Notice that the optimal feedback laws given by (5.17) depend on A, B, R, and Q (partly through dependence on the P_t^o sequence) but are independent of the variance matrix V_t of the white noises ξ_t. Indeed, exactly the same decision rule would be implied if we set $V_t = 0$ for all t, so that there is no randomness in the system. While the noise statistics V_t don't influence the optimal decision rules, they do influence the value of the maximized criterion function through the dependence of d_{t_0} on V_t.

3. The Basic Convergence and Stability Theorems, Which Require Controllability and Reconstructibility

We now proceed to study the behavior of the solution of our problem as we extend the horizon arbitrarily far into the future, or what amounts to the same thing, as we drive the initial period t_0 toward $-\infty$, holding t_1, fixed. We would find the following two characteristics desirable. First, as we drive $t_0 \to -\infty$, we would like P_{t_0} to converge to a constant matrix P which is independent of the given terminal matrix P_{t_1}. This is a desirable feature because it implies via (5.17) that the sequence of optimal control laws $\{F_{t_0}\}$ also converges to a constant as $t_0 \to -\infty$. This has the practical implication that the feedback law $\{F_t\}$ that solves the infinite horizon problem is time invariant, so that $F_t = F$ for all t, and that the resulting closed loop system

$$x_{t+1} = (A - BF)x_t + \xi_{t+1}$$

is time invariant. Our second *desideratum* is, given that it is time invariant, that the closed loop system be stable. This requires that the matrix $(A - BF)$ be stable, that is, have eigenvalues with moduli less than unity.

We shall state and prove several theorems that taken together give a set of conditions that are sufficient to guarantee these two desirable features.

We first recall that as a result of theorem 5.2, the parameters of the optimal feedback laws $\{F_t\}$ and of the value function matrices $\{P_t\}$ are independent of the matrices V_t of the second moments of the noises ξ_t. Thus for purposes of studying the behavior of F_t^0 and P_t^0 as $t \to -\infty$, we can just as well study the nonrandom problem that results when we set $V_t = 0$ for all t. The problem can be stated as follows: to maximize

$$x_{t_0}^T P_{t_0} x_{t_0} = \sum_{t=t_0}^{t_1-1} \left(x_t^T R x_t + u_t^T Q u_t \right) + x_{t_1}^T P_{t_1} x_t \tag{5.19}$$

subject to x_{t_0} given and the law of motion

$$x_{t+1} = Ax_t + Bu_t.$$

Here R is again a negative semidefinite $(n \times n)$ matrix, Q is a negative definite $(k \times k)$ matrix, and P_{t_1} a given $(n \times n)$ negative semidefinite matrix. The maximization is over $F_{t_1-1}, F_{t_1-2}, \cdots, F_{t_0}$ where

$$u_t = -F_t x_t.$$

We shall study the behavior of the solution of this problem when we take the limit as $t_0 \to -\infty$.

We first state the following theorem.

Theorem 5.3: Consider problem (5.19) with terminal value matrix $P_{t_1} = 0$. Assume that the system (A, B) is controllable. Then the optimal P_t calculated from the matrix Riccati difference equation (5.19) converges as $t_0 \to -\infty$.

Proof: The value of problem (5.19) starting from $x_{t_0} = x_0$ is $x_0^T P_{t_0}^o x_0$, where $P_{t_0}^o$ is the solution of the matrix Riccati difference equation (5.18) starting from $P_{t_1} = 0$. Notice that since R and Q are negative semidefinite and negative definite matrices, respectively, we have

$$\begin{aligned} x_0^T P_{t_0+1}^0 x_0 &= \max_{\{u_s\}_{s=t_0+1}^{t_1-1}} \sum_{t=t_0+1}^{t_1-1} \{x_t^T R x_t + u_t^T Q u_t\} \quad , \text{ given } \quad x_{t_0+1} = x_0 \\ &= \max_{\{u_s\}_{s=t_0}^{t_1-2}} \sum_{t=t_0}^{t_1-2} \{x_t^T R x_t + u_t^T Q u_t\} \quad , \text{ given } \quad x_{t_0} = x_0 \\ &\geq \max_{\{u_s\}_{s=t_0}^{t_1-1}} \sum_{t=t_0}^{t_1-1} x_t^T R x_t + u_t^T Q u_t \quad , \text{ given } \quad x_{t_0} = x_0 \\ &= x_0^T P_{t_0} x_0. \end{aligned}$$

In each case the maximization is subject to the law of motion $x_{t+1} = Ax_t + Bu_t$. Thus we have that for all $t_0 \leq t_1 - 1$,

$$x_0^T P_{t_0+1}^o x_0 \geq x_0^T P_{t_0}^0 x_0 \tag{5.20}$$

for all $x_0 \epsilon R^n$. According to (5.20), for any $x_0 \epsilon R^n$, the sequence $x_0^T P_{t_0-i}^0 x_0$ $i = t_0 - t_1 + 1, t_0 - t_1 + 2, \cdots$ decreases monotonically with increases in the index i. Furthermore, since (A, B) is controllable, for every $x_0 \epsilon R^n$, there exists a control sequence that drives x_0 to the origin in n steps. Consider using such a sequence of controls, followed by zero controls thereafter. This set of controls delivers a value of the criterion function

$$\sum_{t=t_0}^{t_1-1} x_t^T R x_t + u_t^T Q u_t$$

starting from $x_{t_0} = x_0$, that provides a lower bound for the values of the problem for any $t_0 \leq t_1 - 1$. It follows that for every $x_0 \epsilon R^n$, $x_0^T P_{t_0-i} x_0$ is monotonically decreasing as i increases, and is bounded below. Therefore, for every $x_0 \epsilon R^n, \lim_{i \to \infty} x_0^T P_{t_0-i} x_0 = \lim_{t_0 \to -\infty} x_0^T P_{t_0} x_0$ exists. Since this limit exists for every $x_0 \epsilon R^n$, it follows that every element of the matrix P_{t_0} converges as $t_0 \to -\infty$. To see this, first set $x_0 = (1\ 0\ 0\ \cdots\ 0)^T$, and notice that $\lim_{t_0 \to -\infty} x_0^T P_{t_0}^0 x_0$ equals the limit as $t_0 \to -\infty$ of the (1, 1) element of $P_{t_0}^o$. Similarly, setting $x_0 = e_i$, where e_i is the i^{th} unit vector, shows that $\lim_{t_0 \to -\infty} x_0^T P_{t_0}^o x_0$ equals the i^{th} diagonal entry of $P_{t_0}^o$. Next, choose $x_0 = (1\ 1\ 0\ 0\ 0\ \cdots\ 0)^T$, to show that the (1, 2) element of $P^o_{t_0}$ converges as $t_0 \to -\infty$. Proceeding with this argument leads to the conclusion that all elements of $P^o_{t_0}$ converge, to a limit P

$$\lim_{t_0 \to -\infty} P^0(t_0) = P^0. \blacksquare$$

We immediately have

Corollary 5.1: Under the conditions of theorem 5.3, the limiting matrix P of the value function is negative semidefinite and satisfies the ***algebraic matrix Riccati equation***

$$P = A^T PA + R - A^T PB(B^T PB + Q)^{-1} B^T PA. \tag{5.21}$$

Proof: Negative semidefiniteness of P follows from the facts that the matrix Riccati difference equation (5.18) maps negative semidefinite P_t into negative semidefinite P_{t-1}, that $P_{t_1} = 0$ is negative semidefinite, and that limits of sequences of negative semidefinite matrices are negative semidefinite. Equation (5.21) follows by taking limits of both sides of the matrix Riccati equation (5.18) as $t \to -\infty$. $\blacksquare$

If iterations on the matrix Riccati difference equation (5.18) from terminal matrix $P_{t_1} = 0$ converge to a negative semidefinite matrix P as $t_0 \to -\infty$, it follows from (5.17) that $\lim_{t_0 \to -\infty} F^o_t$ exists and equals F^o, say.

We desire to study the stability characteristics of the optimal steady-state closed loop system

$$x_{t+1} = (A - BF^o)x_t.$$

In particular, we would like the steady-state optimal closed loop system matrix $(A - BF)$ to be a stable matrix. The following theorem states one useful set of sufficient, though not necessary, conditions for $(A - BF)$ to be stable.

Theorem 5.4: Consider the optimal linear regulator problem with $P_{t_1} = 0$. Let the $(n \times n)$ positive semidefinite matrix $-R$ be expressed as $G^T G$ where G is $(r \times n)$, $r \leq n$, and r is the rank of R (such a decomposition of R always exists by a theorem in linear algebra). Assume that the pair (A, B) is controllable, and that the pair (A, G) is reconstructible. Then the optimal closed loop system matrix $(A - BF)$ is stable.

Proof: From (5.17) and (5.21), the algebraic matrix Riccati equation can be written as

$$P = (A - BF)^T P (A - BF) + R + F^T Q F.$$

Let $D = (A - BF)$, and write the above equation as

$$P = D^T P D + R + F^T Q F. \tag{5.22}$$

The closed loop system whose stability we desire to establish is

$$x_{t+1} = (A - BF) x_t.$$

To establish stability it suffices to show that for any $x_{t_0} = x_0, \lim_{t\to\infty} x_t = 0$. To this end, notice that

$$\begin{aligned} x_{t+1}^T P x_{t+1} - x_t^T P x_t &= x_t^T D^T P D x_t - x_t^T P x_t \\ &= x_t^T (D^T P D - P) x_t. \end{aligned}$$

By (5.22), this equation can be written as

$$x_{t+1}^T P x_{t+1} - x_t^T P x_t = -x_t^T (R + F^T Q F) x_t.$$

This implies that

$$\begin{aligned} x_{t_0+j+1}^T P x_{t_0+j+1} &= x_{t_0}^T P x_{t_0} \\ &\quad - \sum_{i=0}^{j} x_{t_0+i}^T (R + F^T Q F) x_{t_0+i}. \end{aligned}$$

Since the left-hand side is less than or equal to zero because P is negative semidefinite, and since $(R + F^T Q F)$ is negative semidefinite, it follows that

$$x_{t_0+i}^T (R + F^T Q F) x_{t_0+i}$$

approaches zero as $i \to \infty$. Since Q is negative definite and $R = -G^T G$, it follows that

$$\begin{aligned} \lim_{i\to\infty} G x_{t_0+i} &= 0 \\ \lim_{i\to\infty} F x_{t_0+i} &= 0. \end{aligned}$$

Notice that

$$\begin{bmatrix} GA^{n-1} \\ GA^{n-2} \\ \vdots \\ GA \\ G \end{bmatrix} x_t = \begin{bmatrix} G(x_{t+n-1} + \sum_{i=1}^{n-1} A^{i-1} BF x_{t+n-i-1}) \\ G(x_{t+n-2} + \sum_{i=1}^{n-2} A^{i-1} BF x_{t+n-i-2}) \\ \vdots \\ G(x_{t+1} + BF x_t) \\ G x_t \end{bmatrix} \tag{5.23}$$

From our previous results, since $\lim_{t\to\infty} Gx_t = 0$ and $\lim_{t\to\infty} Fx_t = 0$, the right side of (5.23) has a limit of a zero vector as $t \to \infty$. Therefore, the limit of the left hand side is also zero. But by the assumption that the pair (A, G) is reconstructible, the $(n \cdot r \times n)$ matrix on the left side has rank n and therefore has a left inverse. Therefore, the system of equations

$$\begin{bmatrix} GA^{n-1} \\ GA^{n-2} \\ \vdots \\ GA \\ G \end{bmatrix} x = 0$$

has the unique solution $x = 0$. It follows from the fact that the limit of the right side of (5.23) is zero that the limit of x_t as $t \to \infty$ is zero. This proves that $(A - BF)$ is stable. ∎

4. Convergence and Stability Theorems That Only Require Controllability and Detectability

The following theorem shows how the condition that the pair (A, G) is reconstructible can be relaxed and replaced by the assumption that (A, G) is detectable.

Theorem 5.5: Consider the optimal linear regulator problem, to maximize

$$\sum_{t=t_0}^{t_1-1} \left\{ x_t^T R x_t + u_t^T Q u_t \right\} \tag{5.24}$$

subject to x_{t_0} given, and $x_{t+1} = Ax_t + Bu_t$. Here x_t is $(n \times 1)$ and u_t is $(k \times 1)$, while R is a negative semidefinite matrix of rank $r \leq n$, and Q is negative definite. Assume that the pair (A, B) is controllable. Further, let R be represented as $-R = G^T G$ where the matrix G is $r \times n$. Assume that the pair (A, G) is detectable. Then the closed loop system matrix $(A - BF)$ is stable. Further, the feedback law assumes the form

$$F = (F_1' \ 0)U$$

where F_1' is $(k \times m)$ and where U is any nonsingular $(n \times n)$ matrix in which the first m rows of U form a basis for the reconstructible subspace of (A, G). This means that the optimal setting for u_t is a linear combination of basis vectors for the reconstructible subspace.

Proof: Select a nonsingular matrix U whose first m rows form a basis for the reconstructible subspace of the pair (A, G), where m is the dimension of the reconstructible subspace.

Construct the reconstructibility canonical form by defining $x'_t = Ux_t, A' = UAU^{-1}, G' = GU^{-1}, B' = UB$, so that

$$x'_{t+1} = A'x'_t + B'u_t$$
$$y'_t = G'x'_t$$

or

$$\begin{bmatrix} x'_{1t+1} \\ x'_{2t+1} \end{bmatrix} = \begin{bmatrix} A'_{11} & 0 \\ A'_{21} & A'_{22} \end{bmatrix} \begin{bmatrix} x'_{1t} \\ x'_{2t} \end{bmatrix} + B'u_t \tag{5.25}$$

$$y'_t = \begin{bmatrix} G'_1 & 0 \end{bmatrix} \begin{bmatrix} x'_{1t} \\ x'_{2t} \end{bmatrix} \tag{5.26}$$

Here x'_{1t} is $(m \times 1)$, x'_{2t} is $(n-m) \times 1$, G'_1 is $(m \times m)$, A'_{11} is $(m \times m)$. In terms of the transformed variables x'_t, the term $x_t^T R x_t$ in the criterion function (5.24) is

$$\begin{aligned} x_t^T R x_t &= x_t'^T U^{-1^T} R U^{-1} x'_t \\ &= -x_t'^T U^{-1^T} G^T G U^{-1} x'_t. \end{aligned}$$

In constructing the reconstructibility canonical form, it was proved that $GU^{-1} = (G'_1 \ 0)$ where G'_1 is the $(m \times m)$ matrix in (5.26). Thus, the term in the criterion function can be written as

$$\begin{aligned} x_t^T R x_t &= -\, x_{1t}'^T G_1^T G_1 x'_{1t} \\ &= x_{1t}'^T R'_{11} x'_{1t} = x_t'^T \begin{bmatrix} R'_{11} & 0 \\ 0 & 0 \end{bmatrix} x'_t \end{aligned}$$

where $-R'_{11} = G_1^T G_1$ is an $(m \times m)$ matrix. Thus, the optimum problem posed in the statement of the theorem is equivalent with the following problem: to maximize

$$\sum_{t=t_0}^{t_1-1} \left\{ x_t'^T \begin{bmatrix} R'_{11} & 0 \\ 0 & 0 \end{bmatrix} x'_t + u_t^T Q u_t \right\}$$

subject to

$$\begin{bmatrix} x'_{1t+1} \\ x'_{2t+1} \end{bmatrix} = \begin{bmatrix} A'_{11} & 0 \\ A'_{12} & A'_{22} \end{bmatrix} \begin{bmatrix} x'_{1t} \\ x'_{2t} \end{bmatrix} + \begin{bmatrix} B'_1 \\ B'_2 \end{bmatrix} u_t.$$

Since (A', B') is controllable, it follows from theorem 5.4 that the matrix Riccati equation for this problem starting from $P'_{t_1} = 0$ converges. Let us partition the matrix Riccati equation for this problem conformably with the partitioning of x'_t. The result is, where we omit primes from variables for convenience,

(5.27)
$$\begin{bmatrix} P_1 1(t-1) & P_{12}(t-1) \\ P_{21}(t-1) & P_{22}(t-1) \end{bmatrix} = \begin{bmatrix} R_{11} & 0 \\ 0 & 0 \end{bmatrix}$$

$$+ \begin{bmatrix} A_{11}^T P_{11}(t) A_{11} + A_{11}^T P_{12}(t) A_{21} + A_{21} P_{21}(t) A_{11} + A_{21}^T P_{22}(t) A_{12} \,, & A_{11}^T P_{12}(t) A_{22} + A_{12}^T P_{22}(t) A_{22} \\ A_{22}^T P_{21}(t) A_{11} + A_{22}^T P_{22}(t) A_{21} \,, & A_{22}^T P_{22}(t) A_{22} \end{bmatrix}$$

$$- \begin{bmatrix} A_{11}^T P_{11}(t) B_1 + A_{21}^T P_{21}(t) B_1 + A_{11}^T P_{12}(t) B_2 + A_{21}^T P_{22}(t) B_2 \\ A_{22}^T P_{21}(t) B_1 + A_{22} P_{22}(t) B_2 \end{bmatrix} \left(B^T P(t) B + Q \right)^{-1}$$

$$\cdot \begin{bmatrix} A_{11}^T P_{11}(t) B_1 + A_{21}^T P_{21}(t) B_1 + A_{11}^T P_{12}(t) B_2 + A_{21}^T P_{22}(t) B_2 \\ A_{22}^T P_{21}(t) B_1 + A_{22} P_{22}(t) B_2 \end{bmatrix}^T$$

Inspection of (5.27) immediately shows that starting from $P_{t_1} = 0$, the solution is

$$P_{12}(t) = 0$$

$$P_{21}(t) = 0$$

$$P_{22}(t) = 0$$

for all $t \leq t_1$. Substituting these solutions into the difference equation for $P_{11}(t)$ gives

$$P_{11}(t-1) = R_{11} + A_{11}^T P_{11}(t) A_{11} - A_{11}^T P_{11}(t) B_1 (B_1^T P_{11} B_1 + Q)^{-1}$$

$$\left(B_1^T P_{11}(t) A_{11} \right).$$

This is just the matrix Riccati equation for the subsystem defined by the matrices (A_{11}, B_1, Q, R_{11}). Since this system is controllable, this equation is known to converge. The optimum steady state control law is given by (restoring the primes),

$$(F_1', \, F_2') = \left[(B_1^{T'} P_{11}' B_1' + Q)^{-1} B_1^{T'} P_{11}' B_{11'} A', \; 0 \right]$$

where F_1' is $k \times m$, and F_2' is a $k \times (n-m)$ vector. The closed loop system is then

(5.28)
$$\begin{bmatrix} x_{1t+1}' \\ x_{2t+1}' \end{bmatrix} = \begin{bmatrix} A_{11}' - B_1' F_1' & 0 \\ A_{21}' - B_2' F_1' & A_{22}' \end{bmatrix} \begin{bmatrix} x_{1t}' \\ x_{2t}' \end{bmatrix}$$

The subsystem (A_{11}', B_1') is controllable, since (A', B') is controllable. Further, by the construction of the reconstructibility canonical form, the pair (A_{11}', G_1') is reconstructible.

Therefore, by theorem 5.4, the closed loop system matrix $(A'_{11} - B'_1 F'_1)$ is stable. The eigenvalues of the closed loop system $(A' - B'F')$ on the right side of (5.28) are the eigenvalues of $(A'_{11} - B'_1 F'_1)$ and the eigenvalues of A'_{22}. The eigenvalues of A'_{22} are less than unity in modulus by virtue of the detectability of (A, G) and by the construction of the reconstructibility canonical form. Therefore the closed loop system matrix $(A' - B'F')$ is stable. Recall that

$$A' = UAU^{-1}, B' = UB, x'_t = Ux_t.$$

From the optimal control law $u(t) = -F'x'(t)$, we can calculate the control law in terms of feedback on the original state variables, namely, $u_t = -Fx_t = -(F'U)x_t$, so that $(F'U) = F$ or $(F'_1\ 0)U = F$. Notice that $(A' - B'F') = U(A - BF)U^{-1}$, where U is nonsingular. Thus, since the eigenvalues of $(A' - B'F')$ are all less than unity in modulus, the eigenvalues of $(A - BF)$, which equal those of $(A' - B'F')$, are also all less than unity in modulus. ∎

We can now prove the following theorem, which shows that under general conditions, the matrix Riccati equation converges to a limit matrix P that is independent of the terminal matrix P_{t_1}.

Theorem 5.6: Consider the optimal linear regulator problem starting from $P_{t_1} = 0$. Assume that sufficient conditions are satisfied so that iterations on the matrix Riccati equation starting from terminal matrix $P_{t_1} = 0$ converge, and that the associated stationary closed loop system matrix $(A - BF)$ is stable. Then for any negative semidefinite terminal value matrix P_{t_1}, iterations on the matrix Riccati equation converge to the same negative semidefinite matrix P, i.e., the limit point described in theorem 5.3.

Proof: The value of the optimal linear regulator problem with terminal value matrix P_{t_1} is

$$\begin{aligned} x_{t_0}^T P(t_0; P_{t_1}) x_{t_0} = \max_{\{u_t\}_{t=t_0}^{t_1-1}} \sum_{t=t_0}^{t_1-1} \left\{ x_t^T R x_t + u_t^T Q u_t \right\} \\ + x_{t_1}^T P_{t_1} x_{t_1} \end{aligned}$$

where $P(t_0; P_{t_1})$ is the solution of the matrix Riccati equation, (5.18) at t_0 with terminal condition P_{t_1}. Because the matrix P_{t_1} is negative semidefinite, it is true that

$$x_{t_0}^T P(t_0, 0) x_{t_0} \geq x_{t_0}^T P(t_0, P_{t_1}) x_{t_0} \tag{5.29}$$

for every x_{t_0} belonging to R^n. Notice that if $u_t = -Fx_t$, then

$$\begin{aligned}\sum_{t=t_0}^{t_1-1}\left[x_t^T R x_t + u_t^T Q u_t\right] &= \sum_{t=t_0}^{t_1-1}\left[x_t^T R x_t + x_t^T F^T Q F x_t\right] \\ &= \sum_{t=t_0}^{t_1-1} x_t^T [R + F^T Q F] x_t.\end{aligned}$$

Now consider applying to the problem with terminal value matrix P_{t_1} the steady state control law $u_t = -Fx_t$, where $F = \lim_{t_0\to-\infty} F_{t_0}$ is derived from the problem with zero terminal value matrix. Then we have $x_{t+1} = (A - BF)x_t = Dx_t$ where $D \equiv (A - BF)$. Let $W = (R + F^T QF)$. Then since $x_t = D^{(t-t_0)} x_{t_0}$, we have that under this control the criterion function attains the value

$$x_{t_0}^T\left[D^{T(t_1-t_0)}P(t_1)D^{(t_1-t_0)} + \sum_{t=t_0}^{t_1-1} D^{T(t-t_0)} W D^{(t-t_0)}\right]x_{t_0}$$

for any $x_{t_0} \epsilon R^n$. Since $u_t = -Fx_t$ is not necessarily the optimal control law, this together with (5.29) implies the inequalities,

$$\begin{aligned}x_0^T P(t_0, 0) x_0 &\geq x_0^T P(t_0, P_{t_1}) x_0 \\ &\geq x_0^T\left[D^{T(t_1-t_0)} P_{t_1} D^{(t_1-t_0)} + \sum_{t=t_0}^{t_1-1} D^{T(t-t_0)} W D^{(t-t_0)}\right] x_0\end{aligned} \tag{5.30}$$

for every $x_0 \epsilon R^n$. By assumption, we know that

$$\lim_{t_0\to-\infty} P(t_0, 0) = P. \tag{5.31}$$

Further, since the eigenvalues of $D = A - BF$ are less than unity in modulus, we have that $\lim_{t_0\to-\infty} D^{T(t-t_0)} P_{t_1} D^{(t_1-t_0)} = 0$ for every P_{t_1}. Therefore, the limit of the right side of (5.30) is

$$\lim_{t_0\to-\infty} x_0^T\left[D^{T(t-t_0)} P_{t_1} D^{(t-t_0)} + \sum_{t=t_0}^{t_1-1} D^{T(t-t_0)} W D^{(t-t_0)}\right] x_0 \tag{5.32}$$

$$\lim_{t_0\to-\infty} x_0^T\left[\sum_{t=t_0}^{t_1-1} D^{T(t-t_0)} W D^{(t-t_0)}\right] x_0 = x_0^T P x_0$$

for every $x_0 \epsilon R^n$. Together with the inequality (5.30), the limits (5.31) and (5.32) establish that $\lim_{t_0\to-\infty} P(t_0, P_{t_1}) = P$ for every P_{t_1}. ∎

We can now state:

Theorem 5.7: Under the conditions of theorem 5.6, the algebraic matrix Riccati equation

$$P = A^T PA + R - A^T PB(B^T PB + Q)^{-1} B^T PA \tag{5.33}$$

has a unique negative semidefinite solution P^o.

Proof: We know from theorem that $P^o = \lim_{t \to -\infty} P(t, 0)$ is a negative semidefinite solution of (5.33). If $\hat{P}$ were another negative semidefinite solution of (5.33), we would have from the previous theorem that

$$\lim_{t_0 \to -\infty} P(t_0, \hat{P}) = P^o.$$

But $\hat{P}$ solves (5.33), implying that

$$P(t_0, \hat{P}) = \hat{P}.$$

Therefore $\hat{P} = P^o$. ∎

5. Convergence and Stability Theorems That Only Require Stabilizability and Detectability

We now provide theorems that relax the assumption that (A, B) is controllable, and replace it with the assumption that (A, B) is stabilizable. We consider the optimal linear regulator problem, and assume that (A, B) is stabilizable. Without loss of generality, assume that the system is in controllability canonical form, so that

$$\begin{bmatrix} x_{1t+1} \\ x_{2t+1} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix} + \begin{bmatrix} B_1 \\ 0 \end{bmatrix} u(t) \tag{5.34}$$

where x_{1t} is $(m \times 1)$, x_{2t} is $(n-m) \times 1$, A_{11} is $(m \times m)$ and A_{22} is $(n-m) \times (n-m)$, where m is the dimension of the controllable subspace. The eigenvalues of A_{22} are in modulus less than unity, and the pair (A_{11}, B_1) is controllable by virtue of the stabilizability of (A, B). Partition R conformably with x so that

$$R = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix}$$

where R_{11} is $(m \times m)$, and R_{22} is $(n-m) \times (n-m)$, and $R_{21} = R_{12}^T$. Partition the value matrices P_{t-i} conformably with the partitioning of x_t, so that

$$P_t = \begin{bmatrix} P_{11}(t) & P_{12}(t) \\ P_{21}(t) & P_{22}(t) \end{bmatrix}$$

where P_{11} is $(m \times m)$, P_{22} is $(n-m) \times (n-m)$. Then writing out the matrix Riccati difference equation (5.27) in partitioned form gives

$$\begin{aligned} P_{11}(t-1) = {} & A_{11}^T P_{11}(t) A_{11} + R_{11} \\ & - A_{11}^T P_{11}(t) B_1 (Q + B_1^T P_{11}(t) B)^{-1} B_1^T P_{11}(t) A_{11} \end{aligned} \tag{5.35}$$

$$\begin{aligned} P_{12}(t-1) = {} & A_{11}^T P_{11}(t) A_{12} + A_{11}^T P_{12}(t) A_{22} + R_{12} \\ & - A_{11}^T P_{11}(t) B_1 (Q + B_1^T P_{11}(t) B_1)^{-1} (B_1^T P_{11}(t) A_{12} \\ & + B_1^T P_{12}(t) A_{22}) \end{aligned} \tag{5.36}$$

$$\begin{aligned} P_{22}(t-1) = {} & A_{12}^T P_{11}(t) A_{12} + A_{12}^T P_{12}(t) A_{22} + A_{22}^T P_{21}(t) A_{12} \\ & + A_{22}^T P_{22}(t) A_{22} + R_{22} \\ & - (A_{12}^T P_{11}(t) B_1 + A_{22}^T P_{12}(t) B_1)(Q + B_1^T P_{11}(t) B_1)^{-1} \\ & (B_1^T P_{11}(t) A_{12} + B_1^T P_{12}(t) A_{22}) \end{aligned} \tag{5.37}$$

Equation (5.35) is itself the matrix Riccati equation for the optimal linear regulator problem, to maximize

$$\sum_{t=t_0}^{t_1-1} \{x_{1t}^T R_{11} x_{1t} + u_t^T Q u_t\} + x_{1t_1}^T P_{11}(t_1) x_{t_1}$$

subject to

$$x_{1t+1} = A_{11} x_{1t} + B_1 u_t$$

$$x_{1t_0} \quad \text{given.}$$

Since the pair (A_{11}, B_1) is controllable, we know from theorems 5.5 and 5.6 that $\lim_{t_0 \to -\infty} P_{11}$ (t_0) exists and is independent of the negative semidefinite terminal matrix $P_{11}(t_1)$. Now represent the negative semidefinite matrix R_{11} as $-R_{11} = G^T G$ where G is $(r \times m)$ and r is the rank of R_{11}, with $r \le m$. Assume that the pair (A_{11}, G) is *detectable*. Then from theorem 5.5, we know that the stationary closed loop system matrix $(A_{11} - B_1 F_1)$ is a *stable* matrix.

From the recursive structure of the partitioned matrix Riccati equations (5.35), (5.36), (5.37), it follows that the limiting behavior of $P_{12}(t)$ and $P_{22}(t)$ as $t \to -\infty$ is equivalent with the behavior of the pair of equations derived by replacing $P_{11}(t)$ in (5.36) with its limiting value P_{11}, and then, $P_{11}(t)$ and $P_{12}(t)$ in (5.37) with their limiting values in (5.37), if $P_{12}(t)$ has a limit. Making this replacement for (5.36) gives

$$\begin{aligned} P_{12}(t-1) &= A_{11}^T P_{11} A_{12} + A_{11}^T P_{12}(t) A_{22} + R_{12} \\ &\quad - A_{11}^T P_{11} B_1 (Q + B_1^T P_{11} B_1)^{-1} (B_1^1 P_{11} A_{12} + B_1^T P_{12}(t) A_{22}). \end{aligned} \tag{5.38}$$

Upon noting that

$$\left[A_{11}^T - A_{11}^T P_{11} B_1 (Q + B_1^T P_{11} B_1)^{-1} B_1^T\right] = (A_{11} - B_1 F_1)^T$$

equation (5.38) can be written as

$$P_{12}(t-1) = (A_{11}^T - F_1^T B_1^T) P_{12}(t) A_{22} + R_{12} + (A_{11}^T - F_1^T B_1^T) P_{11} A_{12} \tag{5.39}$$

Since $(A_{11} - B_1 F_1)$ and A_{22} are both stable matrices, it follows from (5.39) that $P_{12}(t)$ converges as $t \to -\infty$, and that this limit is independent of the terminal matrix $P_{12}(t_1)$.

Again, the limiting behavior of $P_{22}(t)$ as $t \to -\infty$ is governed by the equation derived by substituting the limiting values of the "forcing function" $P_{11}(t)$ and $P_{12}(t)$ in (5.38). Letting P_{11} and P_{12} be the limiting values of $P_{11}(t)$ and $P_{12}(t)$, these substitutions give

$$\begin{aligned} P_{22}(t-1) &= A_{12}^T P_{11} A_{12} + A_{12}^T P_{12} A_{22} + A_{22}^T P_{21} A_{12} \\ &\quad - (A_{12}^T P_{11} B_1 + A_{22}^T P_{12} B_1)(Q + B_1^T P_{11} B_1)^{-1} (B_1^T P_{11} A_{12} + B_1^T P_{12} A_{22}) \\ &\quad + R_{22} + A_{22}^T P_{22}(t) A_{22}. \end{aligned} \tag{5.40}$$

Since all the terms on the right side of (5.40) are constants except the last, and since A_{22} is a stable matrix, it follows that as $t \to -\infty$, $P_{22}(t)$ converges to a matrix P_{22} that is independent of the negative semidefinite $P_{22}(t_1)$ chosen.

By partitioning the optimal steady state feedback matrix F conformably with the partitioning of x, we obtain

$$F = (F_1 \ F_2),$$

where F_1 is $k \times m$ and F_2 is $k \times (n-m)$. From formula (5.17) in partitioned form, we obtain

$$F_1 = (Q + B_1^T P_{11} B_1)^{-1} B_1^T P_{11} A_{11} \tag{5.41}$$

$$F_2 = (Q + B_1^T P_{11} B_1)^{-1} B_1^T P_{12} A_{22}. \tag{5.42}$$

The optimal closed loop system matrix is

$$(A - BF) = \begin{bmatrix} A_{11} - B_1 F_1 & A_{12} - B_1 F_2 \\ 0 & A_{22} \end{bmatrix}. \tag{5.43}$$

The closed loop system matrix is stable, since $(A_{11} - B_1 F_1)$ and A_{22} have both been shown to be stable under our assumptions, and since the eigenvalues of $(A - BF)$ are the eigenvalues of $(A_{11} - B_1 F_1)$ and the eigenvalues of A_{22}.

We collect these results in the form of the following theorem:

Theorem 5.8: Consider the linear optimal regulator problem where (A, B) is stabilizable. Without loss of generality let (A, B) be in controllability canonical form, so that

$$\begin{bmatrix} x_{1t+1} \\ x_{2t+1} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix} + \begin{pmatrix} B_1 \\ 0 \end{pmatrix} u_t$$

where (A_{11}, B_1) is controllable and A_{22} is a stable matrix. Write the criterion function in the form

$$\sum_{t=t_0}^{t_1-1} \left\{ [x_{1t}^T x_{2t}^T] \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix} + u_t^T Q u_t \right\}$$

$$+ \begin{bmatrix} x_{1t_1} \\ x_{2t_1} \end{bmatrix}^T P_{t_1} \begin{bmatrix} x_{1t_1} \\ x_{2t_1} \end{bmatrix}.$$

Here P_{t_1} and R are negative semidefinite and Q is negative definite. Let the rank of the negative semidefinite matrix R_{11} be $r \leq m$, where m is the dimension of the controllable subspace. Let $-R_{11} = G^T G$ where G is $r \times m$. Assume that the pair (A_{11}, G) is *detectable*. Then

(*i*) Iterations on the matrix Riccati equation (5.18) converge to a unique negative semidefinite matrix that is independent of the terminal matrix P_{t_1}.

(*ii*) The optimal closed loop system matrix

$$(A - BF) = \begin{bmatrix} A_{11} - B_1 F_1 & A_{12} - B_1 F_2 \\ 0 & A_{22} \end{bmatrix}$$

is *stable*.

It should be remarked that under the conditions of theorem XXXXX , the conclusions of theorems XXXX and XXXX both hold. In particular, so long as (A, B) is stabilizable and

(A_{11}, G) is detectable (where $G^T G = -R_{11}$), the algebraic matrix Riccati equation (5.33) has a unique negative semidefinite solution.

The argument leading up to theorem 5.8 actually establishes more than is stated there. In particular, for the convergence results on P_t and the stability of $(A - BF)$, all that is used is that R_{11} and $P_{11}(t_1)$ are negative semidefinite. The above arguments establish convergence of P_t as $t \to -\infty$ and stability of $(A - BF)$ for arbitrary $P_{12}(t_1), P_{22}(t_1), R_{12}$ and R_{22}. Thus, it is not required to assume that R is negative semidefinite. This result is useful, so we summarize it in a theorem.

Theorem 5.9: Consider the optimal linear regulator problem described in theorem 5.8. Assume that (A, B) is stabilizable, $P_{11}(t_1)$ and R_{11} negative semidefinite, and (A_{11}, G) is detectable where $G^T G = -R_{11}$. Otherwise $R_{12}, R_{22}, P_{12}(t_1)$ and $P_{22}(t_1)$ are arbitrary matrixes. Then

(*i*) Iterations on the matrix Riccati equation converge to a unique matrix independent of P_{t_1}. (The limit matrix $\lim_{t \to -\infty} P_t$ is not necessarily negative semidefinite, although $\lim_{t \to -\infty} P_{11}(t)$ is negative semidefinite.)

(*ii*) The optimal stationary closed loop system matrix $(A - BF)$ is *stable*.

(*iii*) Partitioning $F = (F_1 \; F_2)$ conformably with the partitioning of x, F_1 is independent of R_{12} and R_{22}, while F_2 is independent of R_{22}.

Proof: Parts (*i*) and (*ii*) follow from our preceding remarks and the argument leading to theorem 5.8. Part (*iii*) follows directly from inspection of equations (5.41) and (5.42), along with recollection of the recursive structure of (5.35), (5.36), and (5.37). ∎

It is useful to collect the results of the previous theorems in the form of the following summary theorem:

Theorem 5.10: Consider an optimal linear regulator problem of the form, maximize,

$$\sum_{t=t_0}^{t_1-1} \left\{ \begin{bmatrix} x_{0t} \\ x_{3t} \end{bmatrix}^T \begin{bmatrix} R_{00} & R_{03} \\ R_{30} & R_{33} \end{bmatrix} \begin{bmatrix} x_{0t} \\ x_{3t} \end{bmatrix} + u_t^T Q u_t \right\} \tag{5.44}$$

subject to

$$\begin{bmatrix} x_{0t+1} \\ x_{3t+1} \end{bmatrix} = \begin{bmatrix} A_{00} & A_{03} \\ 0 & A_{33} \end{bmatrix} \begin{bmatrix} x_{0t} \\ x_{3t} \end{bmatrix} + \begin{bmatrix} B_0 \\ 0 \end{bmatrix} u_t \tag{5.45}$$

here $x_t = \begin{bmatrix} x_{0t} \\ x_{3t} \end{bmatrix}$ is an $(n \times 1)$ vector, partitioned into an $(m \times 1)$ component x_{0t} and an $(n-m) \times 1$ component x_{3t}, where $1 < m < n$. Assume that the pair (A_{00}, B_0) is *controllable*, so that (5.45) is a controllability canonical form. Assume that (A, B) is stabilizable, so that the eigenvalues of A_{33} are bounded in modulus by unity. Assume that R_{00} is negative semidefinite, that Q is negative definite, but that R_{03}, R_{30}, R_{33} are unrestricted as to definiteness. Let R_{00} be factored according to $R_{00} = -G^T G$ where G is an $r \times m$ matrix, where r is the rank of R_{00}. Assume that the pair (A_{00}, G) is *detectable* but not reconstructible. Under these conditions, the problem can be transformed to one of the form, maximize

$$(5.46) \qquad \sum_{t=t_0}^{t_1-1} \begin{bmatrix} x'_{1t} \\ x'_{2t} \\ x'_{3t} \end{bmatrix} \begin{bmatrix} R'_{11} & 0 & R'_{13} \\ 0 & 0 & R'_{23} \\ R'_{31} & R'_{32} & R'_{33} \end{bmatrix} \begin{bmatrix} x'_{1t} \\ x'_{2t} \\ x'_{3t} \end{bmatrix}$$

subject to

$$(5.47) \qquad \begin{bmatrix} x'_{1t+1} \\ x'_{2t+1} \\ x'_{3t+1} \end{bmatrix} = \begin{bmatrix} A'_{11} & 0 & A'_{13} \\ A'_{12} & A'_{22} & A'_{23} \\ 0 & 0 & A'_{33} \end{bmatrix} \begin{bmatrix} x'_{1t} \\ x'_{2t} \\ x'_{3t} \end{bmatrix} + \begin{bmatrix} B'_1 \\ B'_2 \\ 0 \end{bmatrix} u_t$$

Letting $R'_{11} = -G'^T G'$, the pair $(A'_{11} = G')$ is reconstructible. To achieve this reformulation of the problem, the vector x_{0t} is partitioned into $x_{0t} = \left[\begin{smallmatrix} x_{1t} \\ x_{2t} \end{smallmatrix}\right]$, where x_{1t} is a $p \times 1$ vector , where $0 < p < m$ is the dimension of the reconstructible subspace of (A_{11}, G) and x_{2t} is $(m-p) \times 1$. Form an $m \times m$ nonsingular matrix U, whose first p rows U_1 form a basis for the row space of the reconstructibility matrix for the pair (A_{11}, G). Then define x'_t according to the transformation

$$(5.48) \qquad \begin{bmatrix} x'_{0t} \\ x'_{3t} \end{bmatrix} = \begin{bmatrix} x'_{1t} \\ x'_{2t} \\ x'_{3t} \end{bmatrix} = \begin{bmatrix} U & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} x_{0t} \\ x_{3t} \end{bmatrix}$$

where here I is the $(n-m) \times (n-m)$ identity matrix. Then in (5.46), $\begin{bmatrix} R'_{13} \\ R'_{23} \end{bmatrix} = U^{-1T} R_{03}, (R'_{31} R'_{32}) = R_{30} U^{-1}, \begin{bmatrix} B'_1 \\ B'_2 \end{bmatrix} = U B_0, A'_{13} = U_1 A_{03}, A'_{23} = U_2 A_{03}$, and R'_{11} is a $(p \times p)$ matrix. Starting from a terminal value matrix $P'_{t_1} = 0$, iterations on the Riccati matrix difference equations converge as $t_0 \to -\infty$. The stationary optimal feedback rule is of the form $u_t = -F'_1 x'_{1t}$, while the stationary optimal closed loop system is of the form

$$(5.49) \qquad \begin{bmatrix} x'_{1t+1} \\ x'_{2t+1} \\ x'_{3t+1} \end{bmatrix} = \begin{bmatrix} A'_{11} - B'_1 F'_1 & 0 & A'_{13} - B'_1 F'_3 \\ A'_{12} - B'_2 F'_1 & A'_{22} & -A'_{23} - B'_2 F'_3 \\ 0 & 0 & A'_{33} \end{bmatrix} \begin{bmatrix} x'_{1t} \\ x'_{2t} \\ x'_{3t} \end{bmatrix}$$

We note that the eigenvalues of the optimal closed loop system are the *uncontrollable poles* (i.e. the eigenvalues of A'_{33}), the *unreconstructible poles* (i.e. the eigenvalues of A'_{22}), and the optimally controlled controllable poles (i.e. the eigenvalues of $A'_{11} - B'_1 F'_1$)). The eigenvalues of A'_{33} are less than unity in modulus by assumption. The eigenvalues of $(A'_{11} - B'_1 F'_1)$ are less than unity in modulus because it is both possible and optimal to set them this way. The eigenvalues of $(A'_{22} - B'_2 F'_2)$ can be located arbitrarily in the complex plane, subject to the condition that complex eigenvalues appear in conjugate pairs. It is *possible* to locate the eigenvalues of $(A'_{22} - B'_2 F'_2)$ arbitrarily in the complex plane, because (A_{00}, B_0) is controllable. However, because eigenvalues of A'_{22} are all unreconstructible, it is *desirable* and optimal to set $F'_2 = 0$, and so not to tamper with the unreconstructible eigenvalues.

Proof: We ask the reader to prove this theorem, which involves only a repackaging of our earlier results. The reader should use the state transformation (5.48) and trace through its implications.

6. Examples

The following three examples have structures that illustrate aspects of theorem 5.10.

Example 1. A firm chooses its capital stock to maximize

$$\sum_{t=t_0}^{t_1-1} [f_0 + f_1 k_t - d/2(k_{t+1} - k_t)^2 - J_t k_t]$$

$$f_0, f_1, d > 0$$

subject to k_{t_0} given and J_t obeying the law of motion

$$J_{t+1} = \lambda J_t, \mid \lambda \mid < 1.$$

Here k_t is the stock of capital, and J_t its rental rate at time t. Define the state vector and control vector as

$$x_t = \begin{bmatrix} k_t \\ J_t \\ 1 \end{bmatrix}, u_t = (k_{t+1} - k_t)$$

Then we have that A, B, R, and Q are given by

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & 1 \end{bmatrix}, B = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$R = \begin{bmatrix} 0 & -\frac{1}{2} & f_1/2 \\ -\frac{1}{2} & 0 & 0 \\ f_1/2 & 0 & f_0 \end{bmatrix}, Q = -d/2$$

Letting $A_{00} = 1$, $B_0 = 1$, we see immediately that (A_{00}, B_0) is controllable, and that $G = 0$ where $-G^T G = R_{00} = 0$. Thus the pair $(A_{00}, G) = (0,0)$ is not reconstructible. Further, the pair $(A_{00}, G) = (0,0)$ is already in the reconstructibility canonical form indicated by the theorem, there simply being no part of the state that is both controllable and reconstructible. Theorem (5.10) then implies that the eigenvalues of the optimal stationary closed loop system equal those of A, namely $(1, \lambda, 1)$. (To see how these results can be achieved by classical methods, see Sargent [1987, Chapters IX and XIV]).

Example 2. We now consider a rational expectations equilibrium model of an industry consisting of m identical firms that face demand schedule

$$P_t = A_0 - A_1 Q_t, \quad A_0 A_1 > 0 \tag{5.50}$$

where P_t is price at t, and $Q_t = mq_t$, where q_t is output of the representative firm. Let output be given by the production function

$$q_t = f_1 k_t + f_2 n_t, \quad f_1 > 0, f_2 > 0 \tag{5.51}$$

where k_t and n_t are capital and employment of the representative firm, respectively. The firm maximizes

$$\sum_{t=t_0}^{t_1-1} \{P_t q_t - c(n_{t+1}, n_t, k_{t+1}, k_t, J_t, w_t)\} \tag{5.52}$$

subject to the cost schedule

$$\begin{aligned} c(n_{t+1}, n_t, k_{t+1}, k_t, J_t, w_t) = J_t k_t + w_t n_t + (d/2)(k_{t+1} - k_t)^2 + \\ (e/2)(n_{t+1} - n_t)^2, d > 0, e > 0 \end{aligned} \tag{5.53}$$

with k_t and n_t given at t, and subject to the laws of motion for the rentals on capital J_t and labor w_t,

$$\begin{aligned} J_t &= \mu J_{t-1}, \quad |\mu| < 1 \\ w_t &= \lambda w_{t-1}, \quad |\lambda| < 1 \end{aligned} \tag{5.54}$$

We define market wide stocks of capital and labor as $K_t = mk_t$ and $N_t = mn_t$.

A rational expectations competitive equilibrium is reproduced by a social planning problem which is to maximize

$$\sum_{t=t_0}^{t_1-1}\left\{\left[A_0 - \frac{A_1}{2}\,Q_t\right]Q_t - mc(n_{t+1}, n_t, k_{t+1}, k_t, J_t, w_t)\right\} \tag{5.55}$$

subject to $Q_t = mq_t$ and (5.53) and (5.54) (see e.g., Sargent [1987, Ch. XIV]). We shall proceed to analyze the equilibrium as follows. We shall apply the results of theorem (5.10) to show that the eigenvalue of A_{22} is 1, this being a controllable but unreconstructible pole of the system. We shall argue that K_t and N_t are "borderline unstable" being governed by the unit pole, and shall explore what this means for their behavior. We shall also show that Q_t is asympototically stable.

Substituting (5.51) and (5.53) into (5.55) gives

$$\sum_{t=t_0}^{t_1-1}\left\{A_0 - \frac{A_1}{2}(f_1K_t + f_2N_t)(f_1K_t + f_2N_t) - \frac{d}{2m}(K_{t+1} - K_t)^2\right.$$
$$\left. - \frac{e}{2m}(N_{t+1} - N_t)^2 - J_tK_t - w_tN_t\right\}$$

Writing out this objective function gives

$$\begin{aligned}\sum_{t=t_0}^{t_1-1}&\left\{[A_0f_1K_t + A_0f_2N_t] - \frac{A_1}{2}[f_1^2K_t^2 + f_2^2N_t^2 + 2f_1f_2K_tN_t]\right.\\ &\left.-\frac{d}{2m}(K_{t+1} - K_t)^2 - \frac{e}{2m}(N_{t+1} - N_t)^2 - J_tK_t - w_tN_t\right\}\end{aligned} \tag{5.56}$$

Define the state vector and control vector as

$$x_t = \begin{bmatrix} K_t \\ N_t \\ J_t \\ w_t \\ 1 \end{bmatrix} \qquad u_t = \begin{bmatrix} K_{t+1} - K_t \\ N_{t+1} - N_t \end{bmatrix}$$

The transition equation is

$$\begin{bmatrix} K_{t+1} \\ N_{t+1} \\ J_{t+1} \\ w_{t+1} \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & \mu & 0 & 0 \\ 0 & 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} K_t \\ N_t \\ J_t \\ wt \\ 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} K_{t+1} - K_t \\ N_{t+1} - N_t \end{bmatrix}$$

or

$$x_{t+1} = Ax_t + Bu_t \tag{5.57}$$

The objective function can be written as

$$\sum_{t=t_0}^{t_1-1} (x_t^T R x_t + u_t^T Q u_t)$$

where

$$R = \begin{bmatrix} -\frac{A_1}{2} f_1^2 & -\frac{A_1}{2} f_1 f_2 & -\frac{1}{2} & 0 & -\frac{A_0}{2} f_1 \\ -\frac{A_1}{2} f_1 f_2 & -\frac{A_1}{2} f_2^2 & 0 & -\frac{1}{2} & -\frac{A_0}{2} f_2 \\ -\frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 & 0 & 0 \\ \frac{A_0}{2} f_1 & \frac{A_0}{2} f_2 & 0 & 0 & 0 \end{bmatrix}$$

$$Q = \begin{bmatrix} -\frac{d}{2m} & 0 \\ 0 & -\frac{e}{2m} \end{bmatrix}$$

which is to be maximized subject to (5.57).

We can partition the matrices A, B, R, Q and the vector x_t to deliver a controllability canonical form of the type called for in (5.44) and (5.45) of theorem (5.10). Thus, in terms of theorem (5.10), we set

$$A_{00} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \; B_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$R_{00} = \begin{bmatrix} -\frac{A_1}{2} f_1^2 & -\frac{A_1}{2} f_1 f_2 \\ -\frac{A_1}{2} f_1 f_2 & -\frac{A_1}{2} f_2^2 \end{bmatrix}$$

It is straightforward to verify that the pair (A_{00}, B_0) is controllable, since the rank of B_0 itself is two. Next we need to "factor" R_{00}. Let $G^T = \sqrt{\frac{A_1}{2}}(f_1\, f_2)$. Then we have that $-G^T G = R_{00}$.

Examining the reconstructibility structure of the pair (A_{00}, G), we must calculate the rank of

$$Q = \begin{bmatrix} G \\ GA_{00} \end{bmatrix} = \begin{bmatrix} f_1\sqrt{\frac{A_1}{2}} & f_2\sqrt{\frac{A_1}{2}} \\ f_1\sqrt{\frac{A_1}{2}} & f_2\sqrt{\frac{A_1}{2}} \end{bmatrix},$$

which is evidently unity. Therefore, the pair (A_{00}, G) fails to be reconstructible.

To produce a reconstructibility canonical form for (A_{00}, G), we set $U_1 = (f_1\, f_2), U_2 = (1\ 0)$, letting $U = \binom{U_1}{U_2}$ be our nonsingular transformation matrix. Evidently, U is a nonsingular matrix whose first row is a basis for the row space of Q. We find that

$$U^{-1} = \begin{bmatrix} 0 & 1 \\ 1/f_2 & f_1/f_2 \end{bmatrix}$$

Following the construction of theorem (5.10) we form

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{11} & A_{12} \end{bmatrix} = UA_{00}U^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$GU^{-1} = (A_1/2\ 0), R_{11} = -G^TG = -A_1/2$$

$$\begin{bmatrix} x'_{1t} \\ x'_{2t} \end{bmatrix} = Ux_{0t} = \begin{bmatrix} f_1 & f_2 \\ 1 & 0 \end{bmatrix} \begin{matrix} k_t \\ N_t \end{matrix}, \ UB_0 = \begin{bmatrix} f_1 & f_2 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} B'_1 \\ B'_2 \end{bmatrix}.$$

We note that x_{1t} is output, while x_{2t} is capital stock. Thus, output is in the reconstructible subspace, but the capital stock is not (neither is the stock of labor).

The transformed system corresponding to (5.47) of theorem (5.10) assumes the form

$$\begin{bmatrix} Q_{t+1} \\ K_{t+1} \\ J_{t+1} \\ w_{t+1} \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & \mu & 0 & 0 \\ 0 & 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} Q_t \\ K_t \\ J_t \\ w_t \\ 1 \end{bmatrix}$$

$$+ \begin{bmatrix} f_1 & f_2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} K_{t+1} - K_t \\ N_{t+1} - N_t \end{bmatrix}$$

Letting F'_1 be the (2×1) vector $\begin{bmatrix} F'_{11} \\ F'_{21} \end{bmatrix}$, we have that the optimal closed loop system can be represented as

$$\begin{bmatrix} Q_{t+1} \\ K_{t+1} \\ J_{t+1} \\ w_{t+1} \\ 1 \end{bmatrix} = \begin{bmatrix} 1 - (f_1F'_{11} + f_2F'_{21}) & 0 & & -B'_1F'_3 & \\ -F'_{11} & 1 & & -B'_2F'_3 & \\ 0 & 0 & \mu & 0 & 0 \\ 0 & 0 & 0 & \lambda & 0 \\ 0 & 0 & Q & 0 & 1 \end{bmatrix} \begin{bmatrix} Q_t \\ K_t \\ J_t \\ w_t \\ 1 \end{bmatrix}$$

or

$$\begin{bmatrix} Q_{t+1} \\ K_{t+1} \\ x_{3t+1} \end{bmatrix} = \begin{bmatrix} A'_{11} - B'_1F'_1 & 0 & -B'_1F'_3 \\ -B'_2F'_1 & 1 & -B'_2F'_3 \\ 0 & 0 & A_{33} \end{bmatrix} \begin{bmatrix} Q_t \\ K_t \\ x_{3t} \end{bmatrix}$$

where $x_{3t}^T \equiv (J_t, w_t, 1)$, and where F'_3 is a (2×3) matrix giving the optimal feedforward part of the controls on the uncontrollable states $(J_t, w_t, 1)$. Theorem (5.10) implies that the controls feedback on output, but not on capital and labor separately. It also implies that the eigenvalue $1 - (f_1F'_{11} + f_2F'_{21}) = A'_{11} - B'_1F'_1$ is strictly less than unity in modulus.

The above closed loop system implies a law of motion for Q_t that can be represented as

$$(5.58) \qquad Q_{t+1} - Q_t = -(f_1F'_{11} + f_2F'_{21})Q_t - (f_1F'_{13} + f_2F'_{23})x_{3t}$$

where $F_3' = \begin{bmatrix} F_{13}' \\ F_{23}' \end{bmatrix}$, where F_{13}' and F_{23}' are each (1×3) vectors, and where $-(f_1 F_{11}' + f_2 F_{21}')$ is less than zero by theorem (5.10).

To illustrate the behavior of this system, suppose that x_{3t} is constant over time. Then (5.58) implies that Q_t converges to the stationary value

$$Q = \frac{-(f_1 F_{13} + f_2 F_{23})}{f_1 F_{11}' + f_2 F_{21}} x_3 \tag{5.59}$$

The laws of motion for capital and labor can be represented as

$$K_{t+1} - K_t = -F_{11}' Q_t - F_{13}' x_{3t}$$
$$N_{t+1} - N_t = -F_{21}' Q_t - F_{23}' x_{3t}$$

Substituting the steady state value of Q given by (5.59) into these equations and rearranging gives

$$K_{t+1} - K_t = \frac{f_1(F_{21}' F_{13}' - F_{11}' F_{23}')}{(f_1 F_{11} + f_2 F_{21}')} x_3$$
$$N_{t+1} - N_t = \frac{f_2(F_{11}' F_{23}' - F_{21}' F_{13}')}{(f_1 F_{11}' + f_2 F_{21}')} x_3$$

Only in the singular case in which $F_{11}' F_{23}' - F_{21}' F_{13}' = 0$ do capital and employment converge. In general, each diverges in opposite directions at equal rates, governed by the unit eigenvalue that corresponds to the controllable but unreconstuctible pole.

We now use classical methods to show that the sign of $(F_{11}' F_{23}' - F_{21}' F_{13}')$ equals that of $(J/f_1 - w/f_2)$, where J is the constant value of J_t and w_t the constant value of w_t that is assumed in this experiment. Notice that in this experiment, in general either capital or labor becomes negative in finite time even though the parameters can be selected to guarantee that output Q_t converges to a positive value. We shall use classical methods to help us interpret this outcome. We shall study an infinite time, discounted version of the problem. The social planning problem is to maximize with respect to sequences for (K_t, N_t)

$$\begin{aligned}\sum_{t=0}^{\infty} \beta^t \Big\{ & [A_0 - \frac{A_1}{2}(f_1 K_t + f_2 N_t)][f_1 K_t + f_2 N_t] \\ & - \frac{d}{2}(K_{t+1} - K_t)^2 - \frac{e}{2}(N_{t+1} - N_t)^2 - J_t K_t - w_t N_t \Big\},\end{aligned}$$

given $N_0 K_0$, and given sequences $\{J_t, w_t\}$.

The Euler equations for K and N are

$$\begin{aligned}\beta A_0 - f_1 - \beta J_t - A_1\beta f_1[f_1K_t + f_2N_t] + d\beta K_{t+1} \\ - d(\beta+1)K_t + dK_{t-1} = 0\end{aligned} \tag{5.60}$$

$$\begin{aligned}\beta A_0 f_2 - \beta w_t - A_1\beta f_2[f_1K_t + f_2N_t] + e\beta N_{t+1} \\ - e(\beta+1)N_t + eN_{t-1} = 0\end{aligned} \tag{5.61}$$

This system is a matrix Euler equation in (K_t, N_t) that can be solved using the matrix polynomial factorization methods of chapter 1 or Hansen and Sargent [1981]. However, in effect because of the existence of a nonreconstructible uncontrollable state, the following alternative approach is available for this special problem. Multiply the Euler equation (5.60) for K by ef_1, multiply the Euler equation (5.61) for N by df_2 and add them. After rearranging one obtains

$$\begin{aligned}de\beta Q_{t+1} - (A_1\beta f_1^2 e + A_1\beta f_2^2 d + ed(\beta+1)Q_t - deQ_{t-1} \\ = ef_1J_t + df_2w_t - (\beta A_0 f_1^2 e + \beta A_0 f_2^2 d),\end{aligned} \tag{5.62}$$

which is a univariate Euler equation in $Q_t = (f_1K_t + f_2N_t)$ only.

Let $[de\beta - (A_1\beta f_1^2 e + A_1\beta f_2^2 d + ed(\beta+1)L + deL^2] = de\beta(1-(\lambda\beta)^{-1}L)(1-\lambda L)$, where λ is less than unity in absolute value. That a unique λ less than unity in absolute value exists that satisfies this operator equation follows from Sargent [ch. IX, figure 4]. The solution the the difference equation (5.62) that maximizes social welfare can be represented as

$$Q_{t+1} = \lambda Q_t + c_0 + \frac{c_1 L^{-1}}{1-\lambda\beta L^{-1}}\ \{ef_1J_t + df_2w_t\} \tag{5.63}$$

where c_0 and c_1 are constants. Evidently, from (5.63) output Q is asymptotically stable. In particular, let us assume that $J_t = J$ for all t and that $w_t = w$ for all t. Then output Q converges eventually, since $\lambda < 1$. In the undiscounted ($\beta = 1$) version of this problem, λ precisely equals the pole corresponding to the reconstructible, controllable part of the state in the transformed version of example 2.

To investigate how capital and labor are behaving in the vicinity of a steady state for output, return to the Euler equation for capital, which can be represented as

$$[d\beta L^{-1} - d(1+\beta) + dL]K_t = \beta J_t + A_1\beta f_1 Q_t - \beta A_0 f_1.$$

Factoring the polynomial in L on the left side and solving as in Sargent, [1987, ch. XI; Interpreting], we find

$$(5.64) \qquad (1-L)K_{t+1} = \frac{f_1}{d}\frac{\beta L^{-1}}{1-\beta L^{-1}}[A_0 - \frac{J_t}{f_1} - A_1 Q_t]$$

Following a similar procedure for N, we find that

$$(5.65) \qquad (1-L)N_{t+1} = \frac{f_2}{e}\frac{\beta L^{-1}}{1-\beta L^{-1}}[A_0 - \frac{w_t}{f_2} - A_1 Q_t]$$

Now suppose that $w_t = w$ and $J_t = J$, which we know implies that Q_t converges, say to Q. Then except for the singular case in which $w/f_2 = J/f_1$, capital and labor are both diverging, one toward $+\infty$, the other toward $-\infty$. If $J/f_1 > w/f_2$ then $K_t \to -\infty, N_t \to +\infty$, while if $J/f_1 < w/f_2$, then $K_t \to +\infty, N_t \to -\infty$. The economic interpretation of this situation is straightforward. The firm can hire or sell all of the labor and capital that it wants at the rentals w and J, respectively. The linear technology $Q_t = f_1 K_t + f_2 N_t$ permits firms to use one factor to produce the other. It is only the costs of adjusting capital and labor which prevent the firms from immediately exploiting this opportunity without limit.

Example 3. We now consider a variant of the model of example 2. The model is identical with the previous one, except that the supply of labor to the industry is less than perfectly elastic. In particular, we now assume that

$$w_t = C_0 + C_1 N_t, \ c_0, c_1 > 0.$$

All other aspects of technology, preferences, and competitiveness remain the same. The rational expectations competitive equilibrium now implicitly maximizes social planning criterion

$$(5.66) \qquad \sum_{t=t_0}^{t_1-1}\Big\{[A_0 - \frac{A_1}{2}A_Q]Q_t - \frac{d}{2m}(K_{t+1}-K_t)^2\frac{e}{2m} (N_{t+1}-N_t)^2 - J_t K_t - C_0 N_t - \frac{C_1}{2}N_t^2\Big\},$$

where the term $(C_0 N_t + \frac{C_1}{2}N_t^2)$ is the area under the supply curve for labor to the industry. This social planning criterion is the intertemporal sum of consumer surplus minus net social costs of production.

For this problem, we define the state vector x_t and control u_t as

$$x_t = \begin{bmatrix} K_t \\ N_t \\ J_t \\ 1 \end{bmatrix}, u_t = \begin{bmatrix} K_{t+1} - K_t \\ N_{t+1} - N_t \end{bmatrix}$$

The transition equation is now

$$x_{t+1} = Ax_t + Bu_t$$

where

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \mu & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

The R and Q matrices for the linear regulator problem corresponding to the social planning problem are

$$R = \begin{bmatrix} -\frac{A_1}{2} f_1^2 & -\frac{A_1}{2} f_1 f_2 & -\frac{1}{2} & -\frac{A_0}{2} f_1 \\ -\frac{A_1}{2} f_1 f_2 & -\frac{A_1}{2} f_2^2 \frac{C_1}{2} & 0 & \frac{(A_0 f_2 - C_1)}{2} \\ -\frac{1}{2} & 0 & 0 & 0 \\ -\frac{A_0}{2} f_1 & \frac{(A_0 f_2 - C_2)}{2} & 0 & 0 \end{bmatrix}$$

Here the partitions are again designed to match the partition required by (5.44) and (5.45) of theorem (5.10).

Notice that R_{00} is now of full rank, since $c_1 \neq 0$. A factorization of R_{00} is $-G^T G$ where

$$G = \sqrt{\frac{A_1}{2}} \begin{bmatrix} f_1 & f_2 \\ 0 & C_1/A_1 \end{bmatrix}$$

The reader is invited to verify that (A_{00}, G) is now reconstructible. Thus, our optimization problem is automatically in the form of (5.46)–(5.47) of theorem 5.10 with the understanding that x_{2t} is empty.

Application of theorem 5.10 now implies that the optimal closed loop system has all of its controllable poles placed at values less than unity in modulus, while the uncontrollable eigenvalue μ is less than unity in modulus. Only the uncontrollable unit eigenvalue corresponding to the constant state 1 lies on or outside the unit circle. These facts imply that both capital and labor are asymptotically stable, as is output. If we performed a version of the experiment studied in the last example, setting $J_t = J$ for all t, we would find that both K_t and N_t converge, and that it is possible to select the parameters of the model so that the stationary values of K and N are both positive.

This example exhibits the important technical role played by a positive C_1 in eleminating the unreconstructible, uncontrollable, and unstable pole whose presence causes the divergence of capital and labor in (the constant J and w version of) example 2.

Computer Example: Interrelated Factor Demands With Adjustment Costs

This section reports the output from issuing the MATLAB command "dynfac". The program dynfac.m computes the equilibrium of a linear quadratic industry model with interrelated costs of adjustment for capital and labor. The equilibrium is computed by mapping a fictitious social planning problem into a linear regulator problem. The output from issuing the command "dynfac" follows.

```
dynfac
echo on
clc
```

This program calculates the equilibrium of a two-factor version of Lucas and Prescott's 1971 model of investment under uncertainty. The model is linear quadratic, and constant terms are omitted. The model is a version of one described by Hansen and Sargent in 1981 and Sargent in Macroeconomic Theory, 1987.

The model illustrates a way of modeling dynamically interrelated demands for factors of production, and also illustrates some technical aspects governing the "stability" of solutions of linear optimal control problems. In particular, the first model analyzed below is one in which imposing the "transverality conditions" does not imply stabilizing the system. See Sargent, "Linear Control, Filtering, and Rational Expectations," U. of Minn. manusc., for technical details.

```
pause      %Press a key to continue demonstration
clc
```

There is a single representative firm producing one good with two factors of production, capital k(t) and labor n(t). Demand for output in the industry is given by

$$p(t) = -A1 * Y(t) + u(t)$$

where p(t) is output price at t and Y(t) is industry output,and where A1 ¿ 0 and u(t) is a random shock to demand with autoregressive representation

$$u(t) = lam * u(t-1) + eu(t)$$

where eu(t) is a white noise.

Output of the representative firm, y(t), is given by

$$y(t) = f1 * k(t) + f2 * n(t)$$

where k(t) is capital of the representative firm and n(t) is employment.

```
pause     %Press a key to continue demonstration
cla
```

The firm rents capital and labor at exogenous rental rates of J(t) and w(t), respectively. These rental rates follow the autoregressive processes

$$w(t) = rho * w(t\text{-}1) + ew(t)$$

$$J(t) = rho1 * J(t\text{-}1) + rho2 * J(t\text{-}2) + eJ(t)$$

where ew(t) and eJ(t) are white noises.

```
pause     %Press a key to continue demonstration
cla
```

The rational expectations equilibrium of the industry is a pair of contingency plans for k(t),n(t) that maximize the social welfare function

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{\infty} \{ \text{Consumer surplus at } t - C(t)\}$$

where

$$\text{consumer surplus} = -.5 * A1 * Y(t)^2 + Y(t)*u(t)$$

$$C(t) = J(t)*k(t) + n(t)*w(t) + v(t)'*Q*v(t)$$

where v(t) = [k(t+1) - k(t), n(t+1) - n(t)]' and Q is a (2x2) positive definite matrix of "adjustment" costs.

```
pause     %Press a key to continue demonstration
cla
```

We shall proceed by mapping the optimum problem into the undiscounted linear regulator problem. The state vector is [k(t),n(t),w(t),u(t),J(t),J(t-1)]', which we denote x(t). The control vector is the (2x1) vector v(t) = [k(t+1) - k(t), n(t+1) - n(t)]'. The linear regulator problem is to maximize

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{\infty} \{ x(t)'*R*x(t) + v(t)'*Q*v(t)\}$$

$$\text{subject to } x(t+1) = A * x(t) + B * v(t) + e(t)$$

where e(t) is a vector white noise.

We proceed to set some sample parameters for our problem and to map our problem into the linear regulator by filling out the matrices A,B,Q,R.

```
pause      %Press a key to see the matrix B for our problem.
cla
B=[1 0; 0 1; 0 0; 0 0; 0 0; 0 0]
B =
     1   0
     0   1
     0   0
     0   0
     0   0
     0   0
pause      %Press key to set parameters of the a.r.  processes.
cla
rho=.9
rho =
0.9000
lam=.8
lam =
0.8000
rho1=1.3; rho2=-.4;
pause      %Press a key to set parameters governing costs.
cla
```

```
f1=1;f2=3

f2 =

3

Q=[25 5; 5 10]

Q =

    25     5

     5    10

  Q=-Q;

pause      %Press key to set demand curve parameter

cla

A1=4;
```

NOTE: You can edit this final to set the parameters of the model at whatever values you want.

Now we move on to create the matrix A of the linear regulator.

```
pause      %Press a key to see A.

cla

A=[1 0 0 0 0 0; 0 1 0 0 0 0; 0 0 rho 0 0 0;0 0 0 lam 0 0;...

0 0 0 0 rho1 rho2; 0 0 0 0 1 0]

A =

1.0000         0         0         0         0         0
     0    1.0000         0         0         0         0
     0         0    0.9000         0         0         0
     0         0         0    0.8000         0         0
     0         0         0         0    1.3000   -0.4000
     0         0         0         0    1.0000         0

pause      %Press a key to create R.

cla

R=[-.5*A1*f2^2 -.5*f1*f2*A1 0 .5*f2 -.5 0;...

-.5*A1*f1*f2 -.5*A1*f1^2 -.5 .5*f1 0 0;0 -.5 0 0 0 0;...
```

```
.5*f2 .5*f1 0 0 0 0;-.5 0 0 0 0 0;0 0 0 0 0 0]
R =
```

-18.0000	-6.0000	0	1.5000	-0.5000	0
-6.0000	-2.0000	-0.5000	0.5000	0	0
0	-0.5000	0	0	0	0
1.5000	0.5000	0	0	0	0
-0.5000	0	0	0	0	0
0	0	0	0	0	0

```
pause     %Press a key to solve the social planning problem
cla
[k,S]=double(A',B',R',Q'); %Working, please wait.
Warning:  Matrix is close to singular or badly scaled.
Results may be inaccurate.  RCOND = 6.062365e-017
```

(The warning is related to a unit endogenous eigenvalue that is present in the system. The warning will disappear when we reformulate the system to "cure" the unit eigenvalue below.) The optimal value function for our problem is given by

x(t)'*S'*x(t)

where S is given by

```
S
S =
```

-31.5793	-10.5264	1.4768	2.2866	-1.4450	0.3827
-10.5264	-3.5088	-4.5077	0.7622	1.1850	-0.5391
1.4768	-4.5077	11.2912	-0.0306	-2.7646	1.3113
2.2866	0.7622	-0.0306	0.1678	-0.1058	0.0452
-1.4450	1.1850	-2.7646	-0.1058	0.8342	-0.3898
0.3827	-0.5391	1.3113	0.0452	-0.3898	0.1827

```
state='[k(t),n(t),w(t),u(t),J(t),J(t-1)]';
```

```
pause     %Press a key to see optimal decision rule
cla
```

The optimal decision rule for the social planning problem (i.e., the rational expectations competitive equilibrium decision rules for [k(t+1) - k(t), n(t+1) - n(t)] are

$$v(t) = -F^*x(t)$$

where F is given by

```
F=k'
F =
```

0.5029	0.1676	-0.1547	-0.0291	0.0683	-0.0290
0.2012	0.0671	0.4781	-0.0117	-0.1527	0.0684

```
state
state =
[k(t),n(t),w(t),u(t),J(t),J(t-1)]
pause     %Press a key to see optimal ''closed loop''.
cla
```

The optimal "closed loop" system is

$$x(t+1) = (A - B^*F) * x(t) + e(t+1)$$

where A - B*F = ABF is given by

```
ABK=A-B*k'
ABK =
```

0.4971	-0.1676	0.1547	0.0291	-0.0683	0.0290
-0.2012	0.9329	-0.4781	0.0117	0.1527	-0.0684
0	0	0.9000	0	0	0
0	0	0	0.8000	0	0
0	0	0	0	1.3000	-0.4000
0	0	0	0	1.0000	0

```
state
state =
[k(t),n(t),w(t),u(t),J(t),J(t-1)]
pause     %Press a key to continue
cla
```

Let's look at the eigenvalues of the "feedback part" of ABK, namely, the (2x2) upper left submatrix. First we form this matrix, call it ABK11:

```
ABK11=ABK(1:2,1:2)
ABK11 =
     0.4971    -0.1676
    -0.2012     0.9329
pause     %Press a key to continue demonstration
cla
```

Now calculate the eigenvalues of ABK11:

```
eig(ABK11)
ans =
     0.4300
     1.0000
```

Notice that there is a unit eigenvalue, so that the closed loop system fails to be "stable". This will be so regardless of how you set the parameters - you can convince yourself of this by editing this file and setting alternative parameter values. Can you figure out why the optimal closed loop system is always unstable for this model? What is the economic interpretation?

```
pause     %Press a key to continue demonstration
cla
```

We now alter the problem by changing the specification of the wage for labor. Instead of assuming that it is exogenous to the industry, we now assume that the industry faces an upward sloping supply curve for labor

$$w(t) = S0 + S1 * N(t)$$

where N(t) is total labor supplied to the industry, and S1 > 0. (Since we are ignoring constants in this problem, i.e., working in deviations from means, we set S0 = 0).

The rational expectations competitive equilibrium under this altered specification can be computed by altering the social planning problem so that costs properly account account for producer surplus. In particular, we should add to our previous definition of C(t) the term $.5^*S1^*n(t)^2$. (Remember that there is a single representative firm, so that we can equate n(t) to N(t) in the social planning problem.) So C(t) becomes

$$C(t) = J(t)^*k(t) + w(t)^*n(t) + v(t)'^*Q^*v(t) \\ + .5^*S1^*n(t)^2.$$

```
pause      %Press a key to continue demonstration
cla
```

We now proceed to alter the linear regulator to accommodate the upward sloping supply curve for labor.

The state space must be altered by dropping w(t) as a state variable, and S1*.5 must be substracted from the (2,2) element of the old R matrix, and the A, R, and B matrices must be made conformable with the new state vector. The new state vector is equal to

$$x(t) = [k(t),n(t),u(t),J(t),J(t-1)]'.$$

```
pause      %Press a key to set value of S1
cla
S1=1
S1 =
    1
```

Now we alter the R,A, and B matrices to accommodate the altered social planning problem.

```
pause      %Press a key to see R
```

```
R=[R(:,1:2),R(:,4:6)];
R=[R(1:2,:);R(4:6,:)];
R(2,2)=R(2,2)-.5*S1
R =
```

-18.0000	-6.0000	1.5000	-0.5000	0
-6.0000	-2.5000	0.5000	0	0
1.5000	0.5000	0	0	0
-0.5000	0	0	0	0
0	0	0	0	0

```
state='[k(t),n(t),u(t),J(t),J(t-1)]'
state =
[k(t),n(t),u(t),J(t),J(t-1)]
pause     %Press a key to see A
A=[A(:,1:2),A(:,4:6)];
A=[A(1:2,:);A(4:6,:)]
A =
```

1.0000	0	0	0	0
0	1.0000	0	0	0
0	0	0.8000	0	0
0	0	0	1.3000	-0.4000
0	0	0	1.0000	0

```
pause     %Press a key to see B
B=B(1:5,:)
B =
```

1	0
0	1
0	0
0	0
0	0

```
pause     %Press a key to compute the equilibrium.
cla
[k,s]=double(A',B',R',Q');    %Working, please wait
```

The new state vector is

```
state
state =
[k(t),n(t),u(t),J(t),J(t-1)]
```

The new optimal control law is

```
F=k'
F =
```

0.5231	0.1095	-0.0298	0.0372	-0.0143
0.1383	0.2523	-0.0096	-0.0578	0.0239

```
pause     %Press key to see new closed loop system matrix
cla
```

The optimal closed loop system matrix ABF=A-B*F is

```
ABF=A-B*F
ABF =
```

0.4769	-0.1095	0.0298	-0.0372	0.0143
-0.1383	0.7477	0.0096	0.0578	-0.0239
0	0	0.8000	0	0
0	0	0	1.3000	-0.4000
0	0	0	1.0000	0

The "feedback part" of ABF is

```
ABF11=ABF(1:2,1:2)
ABF11 =
```

0.4769	-0.1095
-0.1383	0.7477

```
pause      %Press a key to see the eigenvalues of ABF11
cla
eig(ABF11)
ans =
    0.4294
    0.7952
```

Can you explain why the eigenvalues are less than 1 now?

HINT: The answer has something to do with the linearity of the production function in n(t) and k(t).

To learn more about the algebraic and economic structure of this example, see Chapter V of Sargent's "Linear Control, Filtering, and Rational Expectations," Unpublished U. of Minn. manuscript.

```
pause      %Press a key to return to menu.
```

This ends the output from the program "dynfac".

7. The Stochastic Optimal Linear Regulator Problem

Suppose now that we return to the optimal linear regulator problem under uncertainty. The following theorem is useful and immediate:

Theorem 5.11: Consider the optimal linear regulator problem, to maximize

$$E_{t_0}\left\{\sum_{t=t_0}^{t_1-1}(x_t^T R x_t + u_t^T Q u_t) + x_{t_1})P_{t_1}x_{t_1}\right\}$$

subject to:

$$x_{t+1} = Ax_t + Bu_t + \xi_{t+1}$$

where ξ_{t+1} is an $(n \times 1)$ vector white noise with $E\xi_t\xi_t^T = V_t$ and where V_t is a positive semidefinite matrix. Assume that (A, B) is *stabilizable*, and without loss of generality that the system is in controllability canonical form. Let R be partitioned conformably with the partitioning of x for the controllability canonical form, and let $-R_{11} = G^T G$ where G is $(r \times m), r \leq m$. Assume that (A, G) is *detectable*. Assume that R_{11} is negative semidefinite and Q negative definite. Consider the criterion,

$$\lim_{t_0 \to -\infty}\left(\frac{1}{t_1 - t_0}\right) E_{t_0} \sum_{t=t_0}^{t_1-1}(x_t^T R x_t + u_t^T Q u_t). \tag{5.67}$$

The optimal steady state control law $u_t = -Fx_t$ maximizes the criterion (5.67), subject to the law of motion $x_{t+1} = Ax_t + Bu_t + \xi_{t+1}$. When $V_t = V$ for all t, the maximal value of the criterion (5.67) is $tr[PV]$ where P is the stationary solution of the matrix Riccati difference equation starting from a negative semidefinite terminal matrix P_{t_1}. This completes the statement of the theorem.

8. The Transofrmation Between Discounted and Undiscounted Problems

We now consider a simple transformation that permits the preceding body of results to apply to discounted problems. Consider the problem to maximize

$$\lim_{t_1 \to -\infty} E_0 \sum_{t=0}^{t_1-1} b^t \left\{x_t^T R x_t + v_t^T Q v_t\right\},\ 0 < b < 1 \tag{5.68}$$

subject to x_0 given and

$$x_{t+1} = Ax_t + Bv_t + \xi_{t+1} \tag{5.69}$$

where $E\xi_t = 0$ for all t and $E\xi_t\xi_t^T = V_t$. Here b is a discount factor which is strictly less than unity in absolute value. For convenience, we have set initial time $t_0 = 0$.

Now define the transformed state variables

$$\bar{x}_t = b^{t/2}x_t$$

$$\bar{v}_t = b^{t/2}v_t$$

so that

$$x_t = b^{-t/2}\bar{x}_t, v_t = b^{-t/2}\bar{v}_t.$$

Substituting these expressions for x_t into the criterion function (5.45), and transition law (5.46) gives the alternative representation of the criterion function

$$(5.70) \qquad \lim_{t_1\to\infty} E_0 \sum_{t=0}^{t_1-1} (\bar{x}_t^T R\bar{x}_t + \bar{v}_t^T Q\bar{v}_t)$$

with the alternative representation of the transition law

$$(5.71) \qquad \bar{x}_{t+1} = b^{\frac{1}{2}} A\bar{x}_t + b^{\frac{1}{2}} B\bar{v}_t + b^{\frac{t+1}{2}} \xi_{t+1}$$

or

$$\bar{x}_{t+1} = \bar{A}\bar{x}_t + \bar{B}\bar{v}_t + b^{\frac{t+1}{2}}\xi_{t+1}, \quad \text{where} \quad \bar{A} = b^{\frac{1}{2}}A, \bar{B} = b^{\frac{1}{2}}B.$$

Now let the eigenvalues of A be $\lambda_1, \lambda_2, \ldots, \lambda_n$. Further, suppose that the original system (5.68) is in controllability canonical form and that the dimension of the controllable subspace is m. So we have that

$$A = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix},$$

where A_{11} is an $m \times m$ matrix. Let the eigenvalues of A_{11} be $\lambda_1, \ldots,, \lambda_m$, and those of A_{22} be $\lambda_{m+1}, \cdots, \lambda_n$.

At this point, the following lemma is handy.

Lemma 5.2: The eigenvalue of $b^{\frac{1}{2}}A$ are $b^{\frac{1}{2}}\lambda_1, b^{\frac{1}{2}}\lambda_2, \cdots, b^{\frac{1}{2}}\lambda_n$.

Proof: From the definitions of the eigenvalues of A and $\bar{A}$. ∎

Now consider the undiscounted problem of maximizing (5.70) subject to (5.71). For convenience, let the rank of R_{11} be m. Suppose that the transformed system (5.70) is stabilizable. Theorem 5.8 guarantees convergence of the matrix Riccati difference equation, the existence of a steady state feedback law $\bar{F}$, and the stability of the closed loop system. The optimal control law of $\bar{v}_t$ is

$$\bar{v}_t = -\bar{F}\bar{x}_t$$

where

$$\bar{F} = (\bar{B}^T\bar{P}\bar{B} + Q)^{-1}\bar{B}^T\bar{P}\bar{A}$$

or

$$\bar{F} = b^2(bB^T\bar{P}B + Q)^{-1}B^T\bar{P}A$$

and where $\bar{P}$ is the solution of the algebraic Riccati equation associated with the transformed system (5.70) and (5.71). In terms of the original variables, we have

$$v_t = -\bar{F}x_t.$$

The closed loop system for the transformed variables is

$$\bar{x}_{t+1} = [b^{\frac{1}{2}}A - b^{\frac{1}{2}}B\bar{F}]\bar{x}_t + b^{\frac{t+1}{2}}\xi_{t+1},$$

which is asymptotically stable by theorem 5.8. The closed loop system in terms of the original variables is then

$$x_{t+1} = [A - B\bar{F}]x_t + \xi_{t+1}. \tag{5.72}$$

Let the eigenvalues of $[A - B\bar{F}]$ be $(\mu_1, \mu_2, \ldots, \mu_n)$. Lemma 5.2 implies that the eigenvalues of $[b^{\frac{1}{2}}A - b^{\frac{1}{2}}B\bar{F}]$ are $b^{\frac{1}{2}}\mu_1, \ldots, b^{\frac{1}{2}}\mu_n$. Theorem 5.8 then implies that $b^{\frac{1}{2}}\mu_1, \ldots, b^{\frac{1}{2}}\mu_n$ are all less than unity in modulus. This in turn implies that

$$|\mu_i| < \frac{1}{\sqrt{b}} \quad i = 1, \ldots, n.$$

Thus the closed loop system (5.72) is "of exponential order less than $\frac{1}{\sqrt{b}}$."

Since the original system was assumed to be in controllability canonical form, so is the transformed system. Writing out the state difference equation, we have

$$\begin{aligned}\begin{bmatrix}\bar{x}_{1t+1}\\ \bar{x}_{2t+1}\end{bmatrix} &= \begin{bmatrix}A_{11}b^{\frac{1}{2}} & A_{12}b^{\frac{1}{2}}\\ 0 & A_{22}b^{\frac{1}{2}}\end{bmatrix}\begin{bmatrix}\bar{x}_{1t}\\ \bar{x}_{2t}\end{bmatrix}\\ &+ \begin{pmatrix}B_1b^{\frac{1}{2}}\\ 0\end{pmatrix}\bar{n}_t + b^{\frac{t+1}{2}}\xi_{t+1}.\end{aligned}$$

The transformed system is stabilizable if the pair (A_{11}, B_1) is controllable, and if the eigenvalues of $A_{22}b^{\frac{1}{2}}$ are all less than unity in absolute value. This last condition is equivalent with the eigenvalues of A_{22} all being less than $\frac{1}{\sqrt{b}}$ in absolute value.

9. Solving the Linear Regulator Problem Via Stochastic Lagrange Multipliers

We return to the nonstochastic optimal linear regulator problem: to maximize

$$\sum_{t=t_0}^{t_1-1} \left\{x_t^T R x_t + u_t^T Q u_t\right\} + x_{t_1}^T P_{t_1} x_{t_1} \quad \text{subject to} \quad x_{t+1} = Ax_t + Bu_t.$$

subject to $x_{t+1} = Ax_t + Bu_t$ where R and P_{t_1} are given negative semi-definite matrices and A is negative definite. We now solve this problem using Lagrange multipliers. This will give rise to the *discrete time maximum principle*.

We form the Lagrangian

$$J = \sum_{t=t_0}^{t_1-1} \left\{x_t^T R x_t + u_t^T Q u_t + 2\lambda_{t+1}^T [Ax_t + Bu_t - x_{t+1}]\right\} + x_{t_1}^T P_{t_1} x_{t_1}. \tag{5.73}$$

Here $\{\lambda_t, t = t_0 + 1, \ldots, t_1\}$ is a sequence of $(n \times 1)$ vectors of Lagrange multipliers. We obtain first-order necessary conditions by differentiating the right side of (5.73) with respect to $\{u_t,\ t = t_0, \ldots, t_1 - 1\}$ with respect to $\{x_t,\ t = t_0 + 1, \ldots, t_1\}$ and equating these derivatives to zero. Differentiating the right side of (5.73) with respect to u_t and equating to zero gives

$$2Qu_t + 2B^T \lambda_{t+1} = 0$$

or

$$u_t = -Q^{-1} B^T \lambda_{t+1}. \tag{5.74}$$

Differentiating with respect to the x_t's and equating to zero gives

$$\lambda_t = Rx_t + A^T \lambda_{t+1}, t = t_0 + 1, \ldots, t_1 - 1 \tag{5.75}$$

$$\lambda_{t_1} = P_{t_1} x_{t_1}, \tag{5.76}$$

Equation (5.75) is called the "co-state equation". Substituting (5.74) into the transition equation $x_{t+1} = Ax_t + Bu_t$ gives

$$x_{t+1} = Ax_t - BQ^{-1}B^T\lambda_{t+1} \tag{5.77}$$

Combining (5.77) and (5.75), we have the homogeneous vector difference equation in the pair (x_t, λ_{t+1})

$$\begin{bmatrix} x_{t+1} \\ \lambda_t \end{bmatrix} = \begin{bmatrix} A & -BQ^{-1}B^T \\ R & A^T \end{bmatrix} \begin{bmatrix} x_t \\ \lambda_{t+1} \end{bmatrix} \tag{5.78}$$

The system (5.78) is to be solved jointly for $(x_t, \lambda_{t+1}; t = t_0, \ldots, t_1)$ subject to the two boundary conditions

$$\begin{aligned} &x_{t_0} \quad \text{given} \\ &\lambda_{t_1} = P_{t_1} x_{t_1} \end{aligned} \tag{5.76}$$

We shall describe two ways to go about solving this system.

a. The Riccati Equation Again

The first method involves *guessing* that a solution can be found of the form

$$\lambda_t = P_t x_t \quad \text{for all} \quad t \le t_1,$$

where P_t's are matrices to be determined. Substituting this guess into the first equation of (5.78) gives

$$x_{t+1} = Ax_t - BQ^{-1}B^T P_{t+1} x_{t+1}$$

or

$$x_{t+1} = (I + BQ^{-1}B^T P_{t+1})^{-1} A x_t$$

Substituting this and $\lambda_t = P_t x_t$ into the second equation of (5.78) gives

$$P_t x_t = R x_t + A^T P_{t+1} (I + BQ^{-1}B^T P_{t+1})^{-1} A x_t,$$

which must hold uniformly in x_t. This requires that the following difference equation in P_t be satisfied

$$P_t = R + A^T P_{t+1} (I + BQ^{-1}B^T P_{t+1})^{-1} A$$

This can also be written

$$P_t = R + A^T(P_{t+1}^{-1} + BQ^{-1}B^T)^{-1}A. \tag{5.79}$$

Equation (5.79) is simply an alternative form of the matrix Riccati difference equation. To see this, first let (a, b, c, d) be matrices with d^{-1} and a^{-1} existing, and recall the formula (see Noble and Daniel [p. 29] or Fortman [ ])

$$(a - bd^{-1}c)^{-1} = a^{-1} - a^{-1}b[d - ca^{-1}b]^{-1}ca^{-1}$$

Use this formula with $a^{-1} = P_{t+1}, b = -B, d = Q, c = B^T$ to get

$$(P_{t+1}^{-1} + BQ^{-1}B^T)^{-1} = P_{t+1} - P_{t+1}B[B^TP_{t+1}B + Q]^{-1}B^TP_{t+1}$$

Substituting this into (5.79) gives

$$P_t = R + A^T\Big[P_{t+1} - P_{t+1}B(B^TP_{t+1}B + Q)^{-1}B^TP_{t+1}\Big]A, \tag{5.80}$$

which is the form of the matrix Riccati difference equation that we have usually utilized.

So with $\lambda_t = P_t x_t$, where P_t obeys (5.80) subject to the terminal condition P_{t_1} given, we have generated a solution to the difference equation system (5.78) that satisfies the terminal condition $\lambda_{t_1} = P_{t_1}x_{t_1}$. Since the initial condition for x_{t_0} is also satisfied, we have produced a solution of our system subject to the appropriate boundary conditions. We have a sufficient number of boundary conditions $(2n)$ to make the solution unique.

b. Vaughan's Method

The second method of solving the system avoids the need to solve the Riccati equation iteratively, but obtains it as a function of the eigenvectors of the state-co-state transition matrix of (5.80). The method is due to Vaughan.[2]

As a preliminary, using the lag operator L, we can write (5.55) as

$$\begin{bmatrix} (L^{-1}I - A) & BQ^{-1}B^T \\ -R & (LI - A) \end{bmatrix} \begin{bmatrix} x_t \\ \lambda_{t+1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{5.81}$$

[2] Vaughan, David R., "A Nonrecursive Algebraic Solution for the Discrete Riccati Equation," *IEEE Transactions in Automatic Control*, October 1970, pp. 597-599.

The dynamic behavior of the system is governed by the zeroes of the characteristic polynomial of the system, namely the solutions of

$$\det \begin{bmatrix} (z^{-1}I - A) & BQ^{-1}B^T \\ -R & (LI - A) \end{bmatrix} = 0 \tag{5.82}$$

Recall the formula for the determinant of a partitioned matrix

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \det d \det(a - bd^{-1}c) = \det a \det (d - ca^{-1}b) \tag{5.83}$$

Applying the identity (5.83) to (5.82), we immediately find that if z_0 is a zero of (5.82), then so is z_0^{-1}. So the zeroes of (5.82) come in reciprocal pairs. Thus the characteristic polynomial in (5.82) has an "Euler-equation like" structure. We shall study this structure further below.

For the infinite horizon problem that emerges when $(t_1 - t_0) \to \infty$, suppose that conditions are met such that there is a stable asymptotic closed loop system matrix $(A - BF)$. Under these circumstances, the optimal solution of (5.81) for the infinite time problem is to solve the "stable roots backwards" and "the unstable roots forwards." Following Vaughan, this insight permits deriving a convenient formula for the limiting value P of the matrix Riccati equation.

To proceed, we follow Vaughan and assume that A is nonsingular. Then rearrange system (5.78) to

$$\begin{pmatrix} x_t \\ \lambda_t \end{pmatrix} = \begin{pmatrix} A^{-1} & A^{-1}BQB^T \\ RA^{-1} & (RA^{-1}BQ^{-1}B^T + A^T) \end{pmatrix} \begin{pmatrix} x_{t+1} \\ \lambda_{t+1} \end{pmatrix} \tag{5.84}$$

Using (5.82), it can be verified that the zeroes of the characteristic polynomial of (5.84) equal those of the characteristic polynomial of (5.81), which must be true because (5.81) and (5.84) describe the same system. These zeroes also equal the eigenvalues of the matrix on the right side of (5.84), call it M, so that we have

$$\begin{pmatrix} x_t \\ \lambda_t \end{pmatrix} = M \begin{pmatrix} x_{t+1} \\ \lambda_{t+1} \end{pmatrix} \tag{5.85}$$

Assume that the eigenvalues of M are distinct. Since the eigenvalues of M come in reciprocal pairs, we can represent M as

$$M = WDW^{-1} \tag{5.86}$$

where

$$D = \begin{pmatrix} \Lambda & 0 \\ 0 & \Lambda^{-1} \end{pmatrix}$$

where Λ is the $(n \times n)$ diagonal matrix whose diagonal elements are the eigenvalues of M that exceed unity in absolute value, and W is the matrix of eigenvectors corresponding to the eigenvalues in D. Inverting, (5.62), and using (5.86), we have

$$\begin{bmatrix} x_{t+1} \\ \lambda_{t+1} \end{bmatrix} = W D^{-1} W^{-1} \begin{bmatrix} x_t \\ \lambda_t \end{bmatrix}$$

The solution of this system is

$$\begin{bmatrix} x_{t+j} \\ \lambda_{t+j} \end{bmatrix} = W \begin{bmatrix} \Lambda^{-j} & 0 \\ 0 & \Lambda^{j} \end{bmatrix} \begin{bmatrix} V_{11} x_t + V_{12} \lambda_t \\ V_{21} x_t + V_{22} \lambda_t \end{bmatrix} \tag{5.87}$$

where

$$W = \begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{bmatrix}, W^{-1} = \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix} \tag{5.88}$$

and W_{ij} and V_{ij} are each $(n \times n)$.

For the infinite time problem, we have already seen that the shadow price λ_t must obey

$$\lambda_t = P x_t \tag{5.89}$$

where P is the limit point of backward iterations on the matrix Riccati difference equation. Substituting (5.89) into (5.87) gives

$$\begin{bmatrix} x_{t+j} \\ \lambda_{t+j} \end{bmatrix} = W \begin{bmatrix} \Lambda^{-j}(V_{11} x_t + V_{12} P x_t) \\ \Lambda^{j}(V_{21} x_t + V_{22} P x_t) \end{bmatrix} \tag{5.90}$$

Under the condition that the optimal closed loop system is stable, we require that $\lim_{j \to -\infty} x_{t+j} = 0$. Since the diagonal elements of Λ exceed unity by construction, this requires that

$$(V_{21} + V_{22} P) x_t = 0, \tag{5.91}$$

which implies

$$P = -V_{22}^{-1} V_{21}. \tag{5.92}$$

Equation (5.92) expresses the limit point P of iterations on the matrix Riccati difference equation in terms of the partitioned inverse of the eigenvector matrix of M. To get an even handier formula, substitute (5.91) into (5.90) and use $\lambda_{t+j} = Px_{t+j}$ to get

$$\begin{bmatrix} x_{t+j} \\ Px_{t+j} \end{bmatrix} = \begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{bmatrix} \begin{bmatrix} \Lambda^{-j}(V_{11}x_t + V_{12}Px_t) \\ 0 \end{bmatrix}.$$

Multiplying the first n equations by P and equating to Px_{t+j} gives

$$PW_{11}\Lambda^{-j}(V_{11} + V_{12}P)x_t = W_{21}\Lambda^{-j}(V_{11} + V_{12}P)x_t,$$

which implies that

$$P = W_{21}W_{11}^{-1} \tag{5.93}$$

This is Vaughan's formula for the limiting value of P in terms of the partitioned matrix of eigenvectors of the state-to-state transition matrix M of (5.85).

c. The Stochastic Version

We can briefly describe the minor modifications of interpretation required to use the above procedures to solve the stochastic optimal linear regulator: to maximize

$$E_{t_0}\left\{\sum_{t=t_0}^{t_1-1}[x_t^T Rx_t + u_t^T Qu_t] + x_{t_1}^T P_{t_1}x_{t_1}\right\}$$

subject to

$$x_{t+1} = Ax_t + Bu_t + \varepsilon(t+1)$$

where $\varepsilon(t)$ is a vector white noise with

$$E\varepsilon(t)\varepsilon(t)^T = V_t > 0,$$

and where E_t is expectation conditioned on x_t. The relevant Lagrangian becomes

$$J = E_{t_0}\left\{\sum_{t=t_0}^{t_1-1}\left[x_t^T Rx_t + u_t^T Qu_t + \lambda_{t+1}^T[Ax_t + Bu_t + \varepsilon_{t+1} - x_{t+1}]\right] + x_{t_1}^T P_{t_1}x_{t_1}\right\}.$$

The first order necessary conditions can be obtained by using the calculus described by Sargent [1987, ch. XIV] to be

$$u_t = -Q^{-1}B^T E_t\lambda_{t+1} \qquad t = t_0, \ldots, t_1 - 1 \tag{5.94}$$

(5.95) $$\lambda_t = Rx_t + A^T E_t \lambda_{t+1}$$

(5.96) $$E_{t_1-1}\lambda_{t_1} = P_{t_1} E_{t_1-1} x_{t_1}$$

From the form of (5.96), it is natural to guess a solution for $E_t\lambda_{t+1}$ of the form $E_t\lambda_{t+1} = P_{t+1}E_t x_{t+1}$ for all $t < t_1$. Using essentially the same mathematics as above, this guess can be verified, and the matrix Riccati difference equation for P_t can be derived.

d. Relationship to "q" Theories of Investment

We can write our solution for λ_t in the form

$$E_t\lambda_{t+1} = P_{t+1}E_t x_{t+1}$$

or

(5.97) $$E_t\lambda_{t+1} = P_{t+1}(A - BF)x_t$$

In the case of an infinite time problem in which P_t converges to P we have

(5.98) $$E_t\lambda_{t+1} = P(A - BF)x_t$$

Substituting (5.98) into (5.94) gives

(5.99) $$u_t = -Q^{-1}B^T P(A - BF)x_t$$

The form of (5.99) and our earlier result that $u_t = -Fx_t$ where F is the asymptotic feedback law (the limit point of F_t) implies the indentity

(5.100) $$F = Q^{-1}B^T P(A - BF)$$

Solving (5.100) for F gives

$$F = (Q + B^T PB)^{-1}B^T PA$$

which is by now a familiar formula. So (5.94) can be viewed as a reinterpretation of our earlier result that $u_t = -Fx_t$.

e. Cross-Products Between States and Controls in the Criterion Function

We now consider the following optimal linear regulator problem: to maximize

$$\sum_{t=t_0}^{t_1-1}\{x_t^TRx_t + 2x_t^TWu_t + u_t^TQu_t\} + x_{t_1}^TP_{t_1}x_{t_1}$$

subject to

$$x_{t+1} = Ax_t + Bu_t.$$

We solve this problem by forming the Lagrangian

$$J = \sum_{t=t_0}^{t_1-1}\{x_t^TRx_t + 2x_t^TWu_t + u_t^TQu_t + \lambda_{t+1}^T[Ax_t + Bu_t - x_{t+1}]\} + x_{t_1}^TP_{t_1}x_{t_1}.$$

Proceeding exactly as above on page ???, we obtain the first-order necessary conditions

$$u_t = -Q^{-1}B^T\lambda_{t+1} - Q^{-1}Wx_t \tag{5.101}$$

$$\lambda_t = A^T\lambda_{t+1} + Rx_t - Wu_t, t = t_0 + 1, \ldots, t_1 - 1 \tag{5.102}$$

$$\lambda_{t_1} = P_{t_1}x_{t_1} \tag{5.103}$$

Substituting (5.101) into the state transition equation and (5.102) and rearranging gives the system

$$\begin{bmatrix} x_{t+1} \\ \lambda_t \end{bmatrix} \begin{bmatrix} A - BQ^{-1}W^T & -BQ^{-1}B^T \\ R - WQ^{-1}W^T & A^T - WQ^{-1}B^T \end{bmatrix} \begin{bmatrix} x_t \\ \lambda_{t+1} \end{bmatrix}. \tag{5.104}$$

Proceeding exactly as above, it is straightforward to show that the zeroes of the characteristic polynomial of the homogeneous difference equation (5.104) come in reciprocal pairs. (Notice the link to the transformation that we described in Chapter 2 to show how to transform a problem with cross-products in states and controls in the objective function into an equivalent problem without cross products.)

As earlier, it can be verified that λ_t obeys

$$\lambda_{t+1} = P_{t+1}x_{t+1}$$

where P_t is the solution of the pertinent matrix Riccati difference equation, in ths case. equation ( ) of Chapter 2. This form of the Riccati equation can be derived using the guess $\lambda_{t+1} = P_{t+1}x_{t+1}$ to solve (5.104), proceeding exactly as above on pp. ???.

For the stochastic optimal linear regulator problem, (5.100), (5.101), (5.102) are replaced with

$$u_t = -Q^{-1}B^T E_t\lambda_{t+1} - Q^{-1}Wx_t$$

$$\lambda_t = A^T E_t\lambda_{t+1} + Rx_t - Wu_t$$

$$E_{t_1-1}\lambda_{t_1} = P_{t_1}E_{t_1-1}x_{t_1}$$

Because of the presence of the term Wx_t, the controls u_t are permitted to be an inexact function of the shadow price $E_t\lambda_{t+1}$.

10. The Inverse Optimal Linear Regulator Problem

We now consider the following problem:

Problem: Given the nonstochastic time invariant system

$$x_{t+1} = Ax_t + Bu_t$$

and the stationary feedback rule

$$u_t = -Fx_t = u_t^*$$

find a return functional of the form

$$J(t_1 - t_0) = x_{t_1}^T P_{t_1} x_{t_1} - \sum_{s=t_0}^{t_1-1} \| G^T x_s + D^T u_s \|^2 \tag{5.105}$$

with $-DD^T = Q < 0$, such that u_t^* is the optimal control law for the asympototic return functional $J = \lim_{t_1-t_0\to\infty} J(t_1 - t_0)$ for *every* $P_{t_1} = P_{t_1}^T \leq 0$, and such that the maximum taken on by J for $u_t = u_t^*$ is independent of the choice of P_{t_1}.

Note that $DD^T = Q$ is nonsingular. The matrix Q can be regarded as given or as chosen arbitrarily. Thus, the problem is, given A, B and the closed loop system $x_{t+1} = (A - BF)x_t$, to find an infinite time optimization problem (i.e., a G and D in (5.82)) such that F is the optimal feedback law for any negative semi-definite terminal value matrix P_{t_1}. Mosca and Zappa[3] formulate this problem and prove the following:

[3] See Edoardo Mosca and Giovanni Zappa, "Consistency Conditions for the Asymptotic Innovations Representation and an Equivalent Inverse Regulation Problem," *IEEE Transactions on Automatic Control*, Vol. AC-24, No. 3, June 1979.

Theorem 5.12: *(Mosca-Zappa)* The inverse optimal regulator problem has a solution if and only if $(A - BF)$ is a stable matrix.

Proof:

(a) First, we show that if $(A-BF)$ is stable, the problem has a solution. Begin by setting $Q < 0$ arbitrarily. Choose $D = D^T = (-Q)^{\frac{1}{2}}$. Then set $G^T = D^T F$. With this choice of G, $\| G^T x_s + D^T u_s \|^2 = \| D^T F x_s - D^T F x_s \|^2 = 0$, so that $J(t_1 - t_0) = x_t^T P_{t_1} x_{t_1}$. Since the closed loop system is asymptotically stable, we have that for all $x_{t_0}, \lim_{(t_1-t_0)\to\infty} x_{t_1} = 0$, which in turn implies that $\lim_{(t_1-t_0)\to\infty} J(t_1 - t_0) = \lim_{(t_1-t_0)\to\infty} x_{t_1} P_{t_1} x_{t_1} = 0$ for any $P_{t_1} \le 0$. Since the return functional (5.82) is nonpositive, we know that a feedback law that achieves $\lim_{(t_1-t_0)\to\infty} J(t_1 - t_0) = 0$ must be optimal for the infinite time problem.

(b) Now we assume that $u_t^* = -F x_t$ maximizes $J = \lim_{(t_1-t_0)\to\infty} J(t_1 - t_0)$ for all $P_{t_1} \le 0$, and that the corresponding maximum of J, call it J^*, is independent of P_{t_1}. Thus

$$J^* = \lim_{(t_1-t_0)\to\infty} \{x_{t_1}^* P_{t_1} x_{t_1}^* - \sum_{s=t_0}^{t_1-1} \| G^T x_s^* + D^T u_s^* \|^2\} = x_{t_0}^T P x_{t_0}$$

where

$$x_{t+1}^* = A x_t^* + B u_t^*, \quad \text{with} \quad x_{t_0}^* = x_{t_0} \quad \text{given},$$

and where P is the limit of the matrix Riccati difference equation, which by assumption exists for all $P_{t_1} \le 0$ and which is independent of P_{t_1}. In particular, choosing $P_{t_1} = 0$ gives

$$\lim_{(t_1-t_0)\to\infty} - \sum_{s=t_0}^{t_1-1} \| G^T x_s^* + D^T u_s^* \|^2 = x_{t_0}^T P x_{t_0}.$$

Therefore, we must have that $\lim_{(t_1-t_0)\to\infty} x_{t_1}^T P_{t_1} x_{t_1}^*$ exists and equals zero for all $P_{t_1} \le 0$. But $\lim_{(t_1-t_0)\to\infty} x_{t_1}^{*T} P_{t_1} x_{t_1}^* = 0$ for all x_{t_0}, implies that $\lim_{(t_1-t_0)\to\infty} x_{t_1}^* = 0$ for all x_{t_0}. Therefore the system is stable, and the closed loop system matrix $(A - BF)$ is a stable matrix. ∎

We reiterate that given an A, B, and F, we can solve the inverse optimal control problem as follows. Pick any $Q < 0$. Set $D = D^T = (-Q)^{\frac{1}{2}}$, and then set $G^T = D^T F$. We note that the solution to the inverse optimal control problem is not unique, since wecan select Q arbitrarily.

a. Two Examples

We ask the reader to solve the following two inverse optimal control problems.

1. The law of motion of capital at t, k_t, is

$$k_{t+1} = k_t + i_t.$$

where i_t is investment during t. The observed closed loop system for k is

$$k_{t+1} = .9k_t.$$

Determine whether the inverse optimal control problem has a solution, and if it does, find one. (Hint: set $A = 1, B = 1, F = .1, Q = 1$ and proceed.) Then formulate the resulting problem as a classical optimization problem (use the calculus of variations) and solve it by factoring the characteristic polynomial of the Euler equation.

2. The law of motion for capital k_t and its rental are

$$\begin{bmatrix} k_{t+1} \\ w_{t+1} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & .8 \end{bmatrix} \begin{bmatrix} k_t \\ w_t \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} i_t.$$

The observed closed loop system for (k_t, w_t) is

$$\begin{bmatrix} k_{t+1} \\ w_t \end{bmatrix} = \begin{bmatrix} .9 & -.1 \\ 0 & .8 \end{bmatrix} \begin{bmatrix} k_t \\ w_t \end{bmatrix}$$

Determine whether the inverse optimal control problem has a solution. If it does, find one.

Exercises

1. Consider the problem of a firm that tries to maximize

$$(1) \qquad E_0 \sum_{t=0}^{\infty} \beta^t \left\{ f_1 n_t - \frac{f_2}{2} n_t^2 - d/2(n_{t+1} - n_t)^2 - w_t n_t \right\}, \; f_1, f_2, d > 0$$

$$0 < \beta < 1$$

subject to (n_t, w_t) given at t, and

$$w_{t+1} = \lambda w_t + \xi_{t+1}, \quad |\lambda| < \frac{1}{\sqrt{\beta}}$$

where ξ_{t+1} is a white noise for w_t. Here n_t is employment of a factor at t, w_t is its rental at t. The firm is imagined to maximize (1) over linear contingency plans of the form

$$n_{t+1} = L(1, n_t, w_t).$$

Assume that ξ_s is orthogonal to w_t for $s > t$.

(a) Formulate the problem as an *undiscounted* optimal linear regulator problem, defining the appropriate state variables, controls, and matrices A, B, Q, R.

(b) Prove that the system is *not* controllable. Find a basis for the controllable subspace. Find a basis for the uncontrollable subspace.

(c) Find a controllability canonical form for the system. Prove that the system is *stabilizable*.

(d) Use our convergence and stability theorem to prove that

(i) Iterations on the matrix Riccati equation converge, and

(ii) the closed loop system matrix $(A - BF)$ for the *original* system has eigenvalues bounded by $\frac{1}{\sqrt{\beta}}$ in modulus.

(e) Write down the matrix Riccati difference equation, and partition it conformably with the partitioning of (A', B') in the controllability canonical form. Write the difference equation for the P_{11} submatrix, and argue that it is itself a matrix Riccati equation. For what problem is it the matrix Riccati difference equation?

(f) Show how the algebraic Riccati equation satisfied by P_{11} (i.e., the equation resulting from taking the limit as $t_0 \to -\infty$ on both sides of the Riccati difference equation) can be solved analytically using the quadratic formula of high school algebra.

2. Consider the problem of a consumer who seeks to maximize

$$(1) \qquad E_0 \sum_{t=0}^{\infty} \beta^t \{u_1 c_t - \frac{u_2}{2} c_t^2\}, \; u_1, \, u_2 > 0 \quad 0 < \beta < 1$$

subject to

$$y_{t+1} = \lambda y_t + \xi_{t+1}, \qquad |\lambda| < \frac{1}{\sqrt{\beta}}$$
$$A_{t+1} = (1 + r)[A_t + y_t - c_t]$$

(y_t, A_t) given at t, (y_0, A_0) given at 0, $(1+r) < \frac{1}{\sqrt{\beta}}$. Here c_t is consumption, y_t is income, A_t is assets, $r > 0$ the interest rate. The random process ξ_t is a white noise that is orthogonal to y_s for $s < t$.

(a) Formulate this problem as an *undiscounted* optimal linear regulator problem, defining the state, control, A, B, Q, R. (The term $\beta^t u, c$ in (1) might cause you a problem. Try using the budget constraint to express c_t in terms of A_{t+1}, A_t, and y_t, and rearrange the sums in A_{t+1} and A_t into a single sum in A_t.)

(b) Prove that the system is *not* controllable.

(c) Find a controllability canonical form. Prove that the system (A, B) is stabilizable. Prove that when written in the controllability canonical form, the pair (A_{11}, G_1) is *detectable* but not *observable*. Here $-R_{11} = G_1^T G_1$ where G_1 is $r \times m$, where $r = \text{rank}\ (R_{11}) \leq m$.

(d) Argue that the matrix Riccati equation converges and that the associated closed loop system matrix $(A - BF)$ for the original system has eigenvalues of modulus bounded by $\frac{1}{\sqrt{\beta}}$.

(e) If $(1 + r) > \sqrt{\beta})^{-1}$, is the pair (A_{11}, G_1) detectable? If $(1 + r) > \sqrt{\beta})^{-1}$, do you think that iterations on the matrix Riccati difference equation will converge?

3. Consider the optimum problem, to maximize

$$\sum_{t=t_0}^{t_1-1} -(n_{t+1} - 2n_t)^2,$$

subject to n_{t_0} given.

(a) Using the dynamic programming algorithm, compute the optimal controls in feedback form, i.e.,

$$n_{t+1} = L_t n_t, \qquad t = t_0, t_0 + 1, \ldots, t_1 - 1.$$

(b) Prove that iterations on the matrix Riccati equation converge as $t_0 \to -\infty$.

(c) Is the asymptotic closed loop system

$$n_{t+1} = (\lim_{s \to -\infty} L_s) n_t$$

stable? If not, what parts of the sufficient conditions for stability from our convergence theorems fail to be met?

4. Consider the optimum problem, to maximize

$$\sum_{t=t_0}^{t_1-1} -.000005\ n_t^2 - (n_{t+1} - 2n_t)^2$$

starting from n_{t_0} given.

(a) Prove that iteration on the matrix Riccati equation converge as $t_0 \to -\infty$.

(b) Write down the algebraic matrix Riccati equation. Argue that the asymptotic optimal closed loop system is approximately $n_{t+1} = \frac{1}{2} n_t$.

(c) Why does such a "small" difference in the objective functions in this problem and the preceding one lead to such a "big" difference in the optimal rules?

5. Consider the following two-player, linear quadratic *dynamic game*. The $(n \times 1)$ state vector x_t evolves according to the transition equation

$$x_{t+1} = A_t x_t + B_{1t} u_{1t} + B_{2t} u_{2t} + \xi_{t+1} \tag{0}$$

where ξ_{t+1} is a vector white noise with $E\xi_t = 0, E\xi_t \xi_t^T = V_t$; u_{jt} is a $(k_j \times 1)$ vector of controls of agent j. Agent 1 maximizes

$$E_{t_0} \sum_{t=t_0}^{t_1-1} \left(x_t^T R_1 x_t + u_{1t}^T Q_1 u_{1t} + u_{2t}^T S_1 u_{2t} \right) \tag{1}$$

where R_1 and S_1 are negative semidefinite, Q_1 is negative definite. Agent 2 maximizes

$$E_{t_0} \sum_{t=t_0}^{t_1-1} \left(x_t^T R_2 x_t + u_{2t}^T Q_2 u_{2t} + u_{1t}^T S_2 u_{1t} \right) \tag{2}$$

where R_2 and S_2 are negative semidefinite and Q_2 is negative definite. We define a *Nash equilibrium* as follows. Agent j is assumed to employ linear control laws

$$u_{jt} = -F_{jt} x_t, \quad t = t_0, \ldots, t_1 - 1$$

where F_{jt} is a $(k_j \times n)$ matrix. Agent i is assumed to know $\{F_{jt}; t = t_0, \ldots, t_1 - 1\}$. Then agent one's problem is to maximize (1) subject ot the known law of motion (0) *and* the known control law $u_{2t} = -F_{2t} x_t$ of agent two. Symmetrically, agent two's problem is to maximize (2) subject to (0) and $u_{1t} = -F_{1t} x_t$. A *Nash equilibrium* is a pair of sequences $\{F_{1t}, F_{2t}; t = t_0, t_0 + 1, \ldots, t_1 - 1\}$ such that $\{F_{1t}\}$ solves agent one's problem, given $\{F_{2t}\}$, and F_{2t} solves agent two's problem, given $\{F_{1t}\}$.

(a) Show how agent one's problem can be written as, maximize

$$E_{t_0} \sum_{t=t_0}^{t_1-1} \left\{ x_t^T (R_1 + F_{2t}^T S_1 F_{2t}) x_t + u_{1t}^T Q_1 u_{1t} \right\}$$

subject to

$$x_{t+1} = (A_t - B_{2t} F_{2t}) x_t + B_{1t} u_{1t} \xi_{1t+1}.$$

Argue that this is a standard optimal linear regulator problem.

(b) Pose agent two's problem as an optimal linear regulator problem.

(c) Prove that the solution of agent one's problem is given by

$$F_{1t} = (B_{1t}^T P_{1t+1} B_{1t} + Q_1)^{-1} B_{1t}^T P_{1t+1}(A_t - B_{2t}F_{2t}) \tag{3}$$

$$t = t_0, t_0 + 1, \ldots, t_1 - 1$$

where P_{1t} is the solution of the following matrix Riccati difference equation, with terminal condition $P_{1t_1} = 0$:

$$\begin{aligned} P_{1t} &= (A_t - B_{2t}F_{2t})^T P_{1t+1}(A_t - B_{2t}F_{2t} + (R_1 + F_{2t})^T S_1 F_{2t}) \\ &\quad - (A_t - B_{2t}F_{2t})^T P_{1t+1} B_{1t}(B_{1t}^T P_{1t+1} B_{1t} + Q_1)^{-1} B_{1t}^T P_{1t+1}(A_t - B_{2t}F_{2t}) \end{aligned} \tag{4}$$

Prove that the solution of agent two's problem is given by

$$F_{2t} = (B_{2t}^T P_{2t+1} B_{2t} + Q_2)^{-1} B_{2t}^T P_{2t+1}(A_t - B_{1t}F_{1t}) \tag{5}$$

where P_{2t} is the solution of the following matrix Riccati difference equation, with terminal condition $P_{1t_1} = 0$:

$$\begin{aligned} P_{2t} &= (A_t - B_{1t}F^T) P_{2t+1}(a_t - B_{1t}F_{1t}) + (R_2 + F_{1t}^T S_2 F_{1t}) \\ &\quad - (A_t - B_{1t}F_{1t})^T P_{2t+1} B_{2t}(B_{2t}^T P_{2t+1} B_{2t} + Q_2)^{-1} B_{2t}^T P_{2t+1}(A_t - B_{1t}F_{1t}). \end{aligned} \tag{6}$$

(d) Describe how the equilibrium sequences $\{F_{1t}, F_{2t}; t = t_0, t_0 + 1, \ldots, t_1 - 1\}$ can be calculated. ***Hint***: use (3), (4), (5), and (6) and "work backwards" from time $t_1 - 1$. Notice that given P_{1t+1} and P_{2t+1}, equations (3) and (4) are a system of $(k_2 \times n) + (k_1 \times n)$ *linear* equations in the $(k_2 \times n) + (k_1 \times n)$ unknowns in the matrices F_{1t} and F_{2t}.

(e) Notice how j's control law F_{jt} is a function of $\{F_{is}, s \geq t, i \neq j\}$. Thus, agent i's choice of $\{F_{it}; t = t_0, \ldots, t_1 - 1\}$ influences agent j's choice of control laws. However, in the Nash equilibrium of this game, each agent is assumed to ignore the influence that his choice exerts on the other agent's choice. In the Nash equilibrium of a *Stackelberg* or *dominant player* game, the timing of moves is so altered relative to the present game that one of the agents called the leader takes into account the influence that his choices exert on the other agent's choices.

Computer Example: A Linear Quadratic Dynamic Game

This section reports the output from the MATLAB program "judd", which computes the Nash feedback equilibrium of a linear quadratic game proposed by Kenneth Judd. The MATLAB program nnash.m is used to compute the equilibrium, as will be seen below. The equilibrium is computed by iterating on a pair of Ricatti equations that is defined by the choice problems of the two agents (firms) in the model.

The output from "judd" follows.

```
judd
echo on
cla
```

This program computes the Nash feedback equilibrium of a linear quadratic dynamic game. Each of two players solves a linear quadratic optimization problem, taking as given and known the sequence of linear feedback rules used by his opponent.

The particular game analyzed is a price-quantity setting game suggested by Ken Judd.

```
pause      %Press a key to continue
cla
```

There are two firms. There is no uncertainty. Relevant variables are defined as follows:

Ii(t) = inventories of firm i at beginning of t.

qi(t) = production of firm i during period t.

pi(t) = price charged by firm i during period t.

Si(t) = sales made by firm i during period t.

Ei(t) = costs of production of firm i during period t.

Ci(t) = costs of carrying inventories for firm i during t.

It is assumed that costs obey

$\mathrm{Ci}(t) = \mathrm{ci}(1) + \mathrm{ci}(2)^{*}\mathrm{Ii}(t) + .5^{*}\ \mathrm{ci}(3)^{*}\mathrm{Ii}(t)^2$

$\mathrm{Ei}(t) = \mathrm{ei}(1) + \mathrm{ei}(2)^{*}\mathrm{qi}(t) + .5^{*}\ \mathrm{ei}(3)^{*}\mathrm{qi}(t)^2$

where ei(j) and ci(j) are constants.

It is assumed that inventories obey the laws of motion

$$\text{Ii}(t+1) = (1 - \text{del}) * \text{Ii}(t) + \text{qi}(t) - \text{Si}(t)$$

```
pause      %Press a key to continue
cla
```

It is assumed that demand is governed by the linear schedule

$$S(t) = d * p(t) + B$$

where S(t) = [S1(t),S2(t)]', d is a (2x2) negative definite matrix, and B is a vector of constants. Firm i is assumed to maximize the undiscounted sum

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T} \{ \text{pi}(t)*\text{Si}(t) - \text{Ei}(t) - \text{Ci}(t)\}$$

by choosing a control law of the form

$$\text{ui}(t) = -\text{Fi} * x(t)$$

where ui(t) = [pi(t),qi(t)]', and the state x(t) is given by x(t)=[I1(t),I2(t), 1].

```
pause      %Press a key to continue
cla
```

Firm i is assumed to solve its control problem taking the (sequence of) control laws uj(t) = -Fj(t)*x(t) as known and given.

The program computes the limiting values of the control laws (F1(t),F2(t)) as the horizon is extended to infinity.

```
pause      %Now the program will set some parameters.  Press a key.
cla
del=.02;
d=[-1 .5; .5 -1]
d =

    -1.0000      0.5000
```

```
    0.5000      -1.0000

B=[25 25]'

B =

    25

    25

pause      %Press a key to set more parameter

cla

c1=[1 -2 1]

c1 =

1 -2 1

c2=[1 -2 1]

c2 =

1 -2 1

e1= [10 10 3]

e1 =

10 10 3

e2= [10 10 3]

e2 =

10 10 3

del1=1-del

del1 =

0.9800

pause      %Press a key to continue

cla
```

Now we'll create the matrices needed to compute the Nash feedback equilibrium. We will proceed by iterating on pairs of "Ricatti" equations. Player 1 has a regulator problem with matrices r1,w1,q1,s1,m1 in the objective function (see the explanation of these quantities when Nash is called shortly) and matrices a,b1,and b2 in the law of motion (again, see the explanation when nnash is called).

```
a=[del1 0 -del1*B(1);0 del1 -del1*B(2);0 0 1]

a =

0.9800         0   -24.5000
     0    0.9800   -24.5000
     0         0     1.0000

b1=del1*[1 -d(1,1); 0 -d(2,1); 0 0]

b1 =

0.9800    0.9800
     0   -0.4900
     0         0

pause     %Press a key to continue

b2=del1*[0 -d(1,2); 1 -d(2,2);0 0]

b2 =

     0   -0.4900
0.9800    0.9800
     0         0

r1=[.5*c1(3) 0 .5*c1(2); 0 0 0;.5*c1(2) 0 c1(1)]

r1 =

 0.5000    0   -1.0000
      0    0         0
-1.0000    0    1.0000

r2=[0 0 0;0 .5*c2(3) .5*c2(2);0 .5*c2(2) c2(1)]

r2 =

0         0         0
0         0   -1.0000
0   -1.0000    1.0000

pause     %Press a key to continue

r1=-r1;r2=-r2;
```

```
q1=[-.5*e1(3) 0; 0 d(1,1)]

q1 =

-1.5000         0
      0   -1.0000

q2=[-.5*e2(3) 0; 0 d(2,2)]

q2 =

-1.5000         0
      0   -1.0000

pause     %Press a key to continue

m1=[0 0; 0 d(1,2)/2]

m1 =

0         0
0    0.2500

m2=m1

m2 =

0         0
0    0.2500

s1=zeros(2);s2=s1;

pause     %Press a key to continue

w1=[0 0;0 0;-.5*e1(2) B(1)/2]

w1 =

      0         0
      0         0
-5.0000   12.5000

w2=[0 0;0 0;-.5*e2(2) B(2)/2]

w2 =

      0         0
      0         0
```

-5.0000   12.5000

```
pause     %Press a key to call nnash to compute equilibrium
nnash
echo on
cla
```

This program computes the limit of a Nash linear quadratic dynamic game

Player i maximizes

Sum {x'*ri*x + 2 x'*wi*ui +ui'*qi*ui + uj'*si*uj + 2 uj'*mi*ui}

subject to the law of motion

x(t+1) = a*x(t) +b1*u1(t)+b2*u2(t)

and a perceived control law uj(t)= -fj*x(t) for the other player

is nxn; b1 is nxk1; b2 is nxk2;

r1 is nxn; r2 is nxn;

q1 is k1xk1; q2 is k2xk2;

s1 is k2xk2; s2 is k1xk1;

w1 is n x k1

w2 is n x k2

m1 is k2 x k1; m2 is k1 x k2;

```
pause     %Press a key to compute the equilibrium
n=length(a);
[x k1]=size(b1);
[x k2]=size(b2);
v1=eye(k1);
v2=eye(k2);
p1=zeros(n);p2=zeros(n);
f1=rand(k1,n);f2=rand(k2,n);
dd=1;tol=.000000000001;
```

```
t1=clock

t1 =

  1.0e+003 *

    1.9880    0.0120    0.0120    0.0050    0.0150    0.0236

jj=0;
while dd>tol;
f10=f1;f20=f2;
g2=(b2'*p2*b2+q2)2;
g1=(b1'*p1*b1+q1)1;
h2=g2*b2'*p2;
h1=g1*b1'*p1;
f1=(v1-(h1*b2+g1*m1')*(h2*b1+g2*m2')) ((h1*a+g1*w1')-...
  (h1*b2+g1*m1')*(h2*a+g2*w2'));
f2=(h2*a+g2*w2')-(h2*b1+g2*m2')*f1;
a2=a-b2*f2;
a1=a-b1*f1;
p1=a2'*p1*a2+r1+f2'*s1*f2-(a2'*p1*b1+w1-f2'*m1)*f1;
p2=a1'*p2*a1+r2+f1'*s2*f1-(a1'*p2*b2+w2-f1'*m2)*f2;
jj=jj+1;
dd=max(abs(f10-f1))+max(abs(f20-f2));
end
t2=clock;et=etime(t2,t1);
pause     %Press a key to see time it took to compute equilibrium
et

et =

    6.8100

f1;
f2;
pause     %Press a key to see number of iterations on Ricatti needed
```

```
jj

jj =

    20

pause      %Press a key to see Firm 1's feedback rule

f1

f1 =
```

0.2437	0.0272	-6.8279
0.3924	0.1397	-37.7341

Firm 2's feedback rule is

```
f2

f2 =
```

0.0272	0.2437	-6.8279
0.1397	0.3924	-37.7341

```
pause   %Press a key to compute closed loop control law

aaa=a-b1*f1-b2*f2

aaa =
```

0.4251	0.0287	0.6810
0.0287	0.4251	0.6810
0	0	1.0000

Recall that the state is x(t)=[I1(t),I2(t),1]' So the equilibrium law of motion is

$$x(t+1) = aaa * x(t)$$

or

$$x(t+1) = (a - b1*F1 - b2*F2) * x(t)$$

```
pause      %Press a key to continue

cla

pause      %Press a key to calculate the optimal stationary values
```

of the inventory levels [I1(t),I2(t)]'.

```
aa=aaa(1:2,1:2);

tf=eye(2)-aa;

tfi=inv(tf);

xbar=tfi*aaa(1:2,3)

xbar =

    1.2469
    1.2469

pause     %press a key to return to menu
```

This terminates the output of judd. You can use the program nnash.m to compute a nash equilibrium for a game of your design.

Chapter 6
The Optimal Observer

1. Introduction

This chapter heavily exploits duality and the theorems of chapter 5 to state convergence and stability theorems for the Kalman filter. The chapter begins with a derivation of the Kalman filter in the style of Luenberger's optimal observer system. This presentation provides interesting perspectives on the Kalman filter. The chapter also describes the "separation principle" of linear optimal control theory, which states how regulation problems with hidden state variables can be solved.

2. The Optimal Observer Problem

We now define an auxiliary system whose behavior is designed to mimic the behavior of another system

Definition 6.1: The system

$$\hat{x}_{t+1} = \hat{A}_t \hat{x}_t + \hat{B}_t u_t + \hat{C}_t y_t \tag{6.1}$$

is a *full order observer* for the system

$$x_{t+1} = A_t x_t + B_t u_t \tag{6.2a}$$

with measurement equation

$$y_t = C_t x_t + E_t u_t \tag{6.2b}$$

if setting $\hat{x}_{t_0} = x_{t_0}$ implies that $\hat{x}_t = x_t$ for all $t \geq t_0$ and for all $u_t, t \geq t_0$.

Theorem 6.1: The system (6.2) is a full order observer for the system (6.1) if and only if

$$\begin{aligned} \hat{A}_t &= A_t - K_t\ C_t \\ \hat{B}_t &= B_t - K_t\ E_t \\ \hat{C}_t &= K_t \end{aligned} \tag{6.3}$$

where $\{K_t, t \geq t_0\}$ is an arbitrary sequence of matrices.

Proof: Substituting (6.2b) into (6.1), gives

$$\hat{x}_{t+1} = \hat{A}_t\hat{x}_t + \hat{B}_t u_t + \hat{C}_t[C_t x_t + E_t u_t].$$

Subtracting (6.2a) from the above equation gives

$$\begin{aligned}\hat{x}_{t+1} - \hat{x}_{t+1} &= \hat{A}_t(\hat{x} - x_t) + [\hat{A}_t + \hat{C}_t C_t - A_t]x_t \\ &\quad + [\hat{B}_t + \hat{C}_t E_t - B_t]u_t\end{aligned} \tag{6.4}$$

Evidently from (6.4), $\hat{x}_{t_0} = x_{t_0}$ implies $\hat{x}_t = x_t$ for $t > t_0$ for all $\{u_t, t > t_0\}$ if and only if $\hat{A}_t = A_t - \hat{C}_t C_t$ and $\hat{B}_t = B_t - \hat{C}_t E_t$. This is true for any arbitrary sequence of matrices $\hat{C}_t = K_t$. ∎

Substituting formulas (6.3) into (6.1) establishes that the full order observer can be represented

$$\hat{x}_{t+1} = A_t\hat{x}_t + B_t u_t + K_t[y_t - C_t\hat{x}_t) - E_t u_t] \tag{6.5}$$

or

$$\hat{x}_{t+1} = A_t\hat{x}_t + B_t u_t + K_t[y_t - \hat{y}_t] \tag{6.6}$$

where $\hat{y}_t$ is the "previously predicted" value for y_t,

$$\hat{y}_t = C_t\hat{x}_t + E_t u_t.$$

The sense in which $\hat{y}_t$ is the previously predicted value for the measurements y_t will become clear shortly. So (6.5) or (6.6) expresses the "prediction" $\hat{x}_{t+1}$ for x_{t+1} as a function of the "lagged prediction," the control, and the error just realized in predicting the observable variables y_t.

Define the *reconstruction error* in estimating the state as $x_t - \hat{x}_t = e_t$. We can then state the following theorem.

Theorem 6.2: Consider the full order observer for the nonstochastic system (6.2a)-(6.2b). The reconstruction error $e_t = x_t - \hat{x}_t$ satisfies the difference equation

$$e_{t+1} = [A_t - K_t C_t]e_t \quad \text{for} \quad t \geq t_0.$$

Proof: Subtract the state difference equation (6.2a) from (6.6) to get

$$x_{t+1} - \hat{x}_{t+1} = [A_t - K_t \; C_t][x_t - \hat{x}_t]. \; \blacksquare$$

If the reconstruction error in this nonstochastic system has the property that $e_t \to 0$ as $t \to \infty$ for all initial errors e_{t_0}, the full order observer is said to be *asymptotically stable*. Notice that the asymptotic stability of the observer depends on the behavior of the matrices $[A_t - K_t C_t]$ as t gets large. For the case in which A_t and C_t are time invariant, we shall presently study the limiting behavior of this matrix.

The following simple lemma is useful in our study of the stochastic linear observer problem.

Lemma 6.1: Consider the system

$$x_{t+1} = A_t x_t + B_t w_{t+1}, \quad t \geq t_0$$

where w_t is a white noise vector with

$$Ew_t = 0$$

$$Ew_t w_t^T = V_t.$$

Define the mean vector and covariance matrix of x_t, $\quad t > t_0$

$$m_t = Ex_t$$

$$\Sigma_t = E(x_t - m_t)(x_t - m_t)^T.$$

Let x_{t_0} be a random variable with given mean vector m_{t_0} and covariance matrix $\sum_{t_0}$. Assume that x_{t_0} is uncorrelated with w_t for $t > t_0$. Then

$$Ex_{t_0+i} = \Psi(t_0 + i, t_0) m_{t_0} \tag{6.7}$$

where $\Psi(t, t_0)$ is the transition matrix

$$\Psi(t, t_0) = \begin{cases} A_{t-1} A_{t-2} \ldots A_{t_0} & t > t_0 \\ I & t = t_0. \end{cases}$$

Further, Σ_t is the solution of the difference equation

$$\Sigma_{t+1} = A_t \Sigma_t A_t^T + B_t V_{t+1} B_t^T \tag{6.8}$$

with initial condition Σ_{t_0} given.

Proof: Equality (6.7) follows from taking the mathematical expectation of each side of the solution of the state equation

Equality (6.8) follows from writing

$$x_{t+1} - m_{t+1} = A_t x_t + B_t w_{t+1} - A_t E x_t$$

or

$$x_{t+1} - m_{t+1} = A_t x_t + E x_t + B_t w_{t+1}.$$

Multiplying each side of this equality by its tranpose and taking mathematical expectations implies equality (6.9). ∎

The preceding discussion is extended in a straightforward manner to cover the case of a stochastic system.

Definition 6.2: Consider the stochastic system

$$(6.9a) \qquad x_{t+1} = A_t x_t + B_t u_t + w_{1t+1}$$

$$(6.9b) \qquad y_t = C_t x_t + E_t u_t + w_{2t}$$

where w_{1t+1} and w_{2t} are vector white noise random errors satisfying $Ew_{1t+1} = 0, Ew_{2t} = 0$

$$E\begin{bmatrix} w_{1t+1} \\ w_{2t} \end{bmatrix}\begin{bmatrix} w_{1t+1} \\ w_{2t} \end{bmatrix}^T = \begin{bmatrix} V_{1t} & 0 \\ 0 & V_{2t} \end{bmatrix},$$

and $Ew_{1t}w_{2s}^T = 0$ for all t and s. Let the system start at time t_0, and let x_{t_0} be a random variable with mean vector $\hat{x}_{t_0}$ and covariance matrix Σ_0. Consider the auxiliary system

$$(6.10) \qquad \hat{x}_{t+1} = \hat{A}_t \hat{x}_t + \hat{B}_t u_t + \hat{C}_t y_t.$$

The system (6.10) is said to be a full order observer for the system (6.9) if setting $\hat{x}_{t_0} = Ex_{t_0}$ implies that $\hat{x}_t = Ex_t$ for all $t > t_0$ and for all $u_t, t \geq t_0$.

We immediately have the following theorem, whose proof mimics the proof of theorem 6.1.

Theorem 6.3: The system (6.10) is a full-order observer for the system (6.9) if and only if

$$\hat{A}_t = A_t - K_t C_t$$

$$\hat{B}_t = B_t - K_t E_t$$

$$\hat{C}_t = K_t$$

where $\{K_t, t \geq t_0\}$ is an arbitrary sequence of matrices.

We leave the proof as an exercise.

We now consider the *stochastic linear optimal observer problem* whose solution leads us to a version of the celebrated Kalman filter.

Definition 6.3: Consider the discrete time system

(6.9*a*) $$x_{t+1} = Ax_t + Bu_t + w_{1t+1} \qquad t \geq t_0$$

(6.9*b*) $$y_t = Cx_t + Eu_t + w_{2t}$$

where $w_t = \begin{bmatrix} w_{1t+1} \\ w_{2t} \end{bmatrix}$ is a serially uncorrelated random process with mean zero and contemporaneous covariance matrix

$$E \begin{bmatrix} w_{1t+1} \\ w_{2t} \end{bmatrix} [w_{1t+1}^T w_{2t}^T] = \begin{bmatrix} V_{1t} & 0 \\ 0 & V_{2t} \end{bmatrix}.$$

We assume that $Ew_{1t+1}w_{2t+s}^T = 0$ for all s, so that w_1 and w_2 are orthogonal at all leads and lags. In (6.9*a*) and (6.9*b*), x_t is an underlying state vector that is not directly observed, u_t is a vector of controls, y_t is a vector of variables that is directly observed, w_{1t+1} is the error process driving the "hidden" state variables, and w_{2t} is a process of "measurement errors." We assume that x_{t_0} is a random vector with

$$Ex_{t_0} = \bar{x}_0$$

$$E(x_{t_0} - \bar{x}_0)(x_{t_0} - \bar{x}_0)^T = \Sigma_0.$$

We assume that x_{t_0} is orthogonal to w_{t+s} for all $s \geq 0$. Consider the *observer system*

(6.11) $$\hat{x}_{t+1} = A\hat{x}_t + Bu_t + K_t[y_t - C\hat{x}_t - Eu_t].$$

Let W_{t_1} be a given positive definite "weighting" matrix, and let $t_1 \geq t_0$ be fixed. The *stochastic linear optimal observer problem* is to find a sequence of matrices $\{K_{t_0}, K_{t_0+1}, \ldots, K_{t_1-1}\}$ and an initial condition $\hat{x}_{t_0}$ that minimizes

$$E\ \{e_{t_1}^T W_{t_1} e_{t_1}\}$$

where $e_t = x_t - \hat{x}_t$ is the reconstruction error at time t.

Definition 6.4: If $V_{2t} > 0$ for all $t \geq t_0$, the problem is called *nonsingular*.

We shall restrict ourselves at this point to considering the nonsingular observer problem. We further restrict ourselves to the time invariant or homoskedastic case in which $V_{2t} = V_2$ is independent of time.

We proceed to solve the optimal observer problem. Subtracting the observer equation (6.11) from the state equation (6.9a) gives

$$x_{t+1} - \hat{x}_{t+1} = [A - K_t C](x_t - \hat{x}_t + w_{1t+1} - K_t w_{2t})$$

or

$$e_{t+1} = [A - K_t C]e_t + w_{1t+1} - K_t w_{2t}. \tag{6.12}$$

Let $\sum_t$ be the covariance matrix of e_t and let $\bar{e}_t$ be the mean of e_t. Then

$$E\ e_t e_t^T = \Sigma_t \ + \ \bar{e}_t \bar{e}_t^T.$$

Further, we have

$$\begin{aligned} E\ \{e_t^T W_t e_t\} &= E\ \left\{(e_t - \bar{e}_t)^T W_t (e_t - \bar{e}_t)\right\} \\ &\quad + E\ \{\bar{e}_t^T W_t \bar{e}_t\} \\ &= tr\, W_t E\ \left\{(e_t - \bar{e}_t)(e_t - \bar{e}_t)^T\right\} \\ &\quad + \bar{e}_t^T W_t \bar{e}_t\,. \end{aligned}$$

It follows that

$$E\ e_t^T W_t e_t = tr\, W_t \Sigma_t \ + \ \bar{e}_t^T W_t \bar{e}_t. \tag{6.13}$$

Applying Lemma (6.1) to the difference equation (6.12) for e_t for an arbitrary sequence $\{K_t,\, t = t_0, \ldots, t_1 - 1\}$ gives

$$\bar{e}_t = \Psi\,(t, t_0)\, \bar{e}_{t_0}. \tag{6.14}$$

and

$$\Sigma_{t+1} = [A - K_t C]\Sigma_t [A - K_t C]^T + V_1 + K_t V_2 K_t^T \tag{6.15}$$

where the transition matrix $\Psi(t, t_0)$ is given by

$$\Psi(t, t_0) = \begin{cases} (A - K_{t-1}C)(A - K_{t-2}C) \cdots (A - K_{t_0}C) & t > t_0 \\ I & t = t_0 \end{cases}$$

First note that for given K_t, (6.15) implies that Σ_{t+1} is a monotonically increasing function of Σ_t. Next notice that Σ_{t_0} is independent of the choice of the initial conditions $\hat{x}_{t_0}$. To see this, write

$$\begin{aligned} \Sigma_{t_0} &= E\ (e_{t_0} - \bar{e}_{t_0})(e_{t_0} - \bar{e}_{t_0})^T \\ &= E\ \left[(x_{t_0} - \hat{x}_{t_0}) - E\ (x_{t_0} - \hat{x}_{t_0})\right]\left[(x_{t_0} - \bar{x}_{t_0}) - E\ (x_{t_0} - \bar{x}_{t_0})\right]^T \\ &= E\ [x_{t_0} - \bar{x}_0][x_{t_0} - \bar{x}_0]^T = \Sigma_0. \end{aligned}$$

Thus, we have that $\Sigma_{t_0} = \Sigma_0$ independent of the choice of $\hat{x}_{t_0}$. Clearly, since from (6.14) $\bar{e}_t = \Phi(t, t_0)\bar{e}_{t_0}$, the term $\bar{e}^{(T)} W_t \bar{e}_t$ is minimized for any positive definite $\bar{W}_t$ by choosing $\bar{e}_{t_0} = 0$, from (6.14) this choice of $\bar{e}_{t_0}$ implies that $\bar{e}_t = 0$ for all $t \geq t_0$. Setting $\bar{e}_{t_0} = 0$ is accomplished by setting

$$\hat{x}_{t_0} = \bar{x}_0. \tag{6.16}$$

Further, since Σ_t is given by the solution of the difference equation (6.15) starting from initial condition $\Sigma_{t_0} = \Sigma_0$, and since Σ_{t_0} is independent of the choice of $\hat{x}_{t_0}$, if follows from (6.13) that setting $\hat{x}_{t_0} = \bar{x}_0$ is the choice that minimizes $E\ e_t^T W_t e_t = tr\ W_t \Sigma_t + \bar{e}_t^T W_t \bar{e}_t$.

It follows that our problem is reduced to that of minimizing the first term of (6.13) for some $t = t_1 > t_0$. We must choose a sequence of matrices $\{K_{t_0}, K_{t_0+1}, \ldots, K_{t_1-1}\}$ to minimize $tr\ W_{t_1}\Sigma_{t_1}$, where Σ_{t_1} solves

$$\Sigma_{t+1} = [A - K_tC]\Sigma_t[A - K_tC]^T + V_1 + K_t V_2 K_t^T \tag{6.15}$$

with initial condition

$$\Sigma_{t_0} = \Sigma_0$$

given. Evidently, this is equivalent with minimizing Σ_{t_1} with respect to $\{K_t, t = t_0, \ldots, t_1-1\}$ subject to (6.15) and the initial condition $\Sigma_{t_0} = \Sigma_0$. In solving this problem, the following theorem is useful.

Theorem 6.4: Consider the difference equation

$$P_t = \{A - BF_t\}^T P_{t+1}\{A - BF_t\} + R + F_t^T Q F_t \tag{6.17}$$
$$t = t_0, t_0+1, \dots, t_1 - 1, t_1 > t_0$$

with terminal condition $P_{t_1} = P_1$ and where $\{F_t, t = t_0, \dots, t_1 - 1\}$ is an arbitrary sequence of matrices. Let $\bar{P}_t$ be the solution of this difference equation with boundary condition $\bar{P}_{t_1} = P_1$. Consider the difference equation

$$G(s) = [A - BH_{s-1}]^T G_{s-1}[A - BH_{s-1}] + R + H_{s-1}^T Q H_{s-1} \tag{6.18}$$
$$s = t_0+1, t_0+2, \dots, t_1$$

subject to the boundary condition $G_{t_0} = P_1$, and where $H_{s-1} = F_{(t_1+t_0-s)}$. Then the solution of the difference equation (6.18) is

$$G_s = \bar{P}_{(t_1+t_0-s)}.$$

Proof: Define $t_1 + t_0 = t^*$, and $s = (t_1 + t_0) - t$. Note that $t = t_1$ implies $s = t_0$ and $t = t_0$ and $t = t_0$ implies $s = t_1$. Then note that equation (6.17) can be written

$$\begin{aligned} P_{t_0-s} &= [A - BF_{t^*-s}]^T P_{(t^*-(s-1))}(A - BF_{(t^*-s)}) \\ &\quad + R + F_{(t^*-s)}^T Q F_{t^*-s} \end{aligned} \tag{6.19}$$
$$s = t_0+1, t_0+2, \dots, t_1$$

where the boundary condition is $P_{t^*-t_0} = P_1$. Define $G(s) = P_{t_0-s}$ and $H_{s-1} = F_{t_0-s}$. Then (6.18) can be written

$$G_s = [A - BH_{s-1}]^T G_{s-1}[A - BH_{s-1}] + R + H_{s-1}^T Q H_{s-1} \tag{6.20}$$
$$s = t_0+1, t_0+2, \dots, t_1$$

where the boundary condition is now $G_{t_0} = P_1$. Therefore, if $\bar{P}_t, t = t_0, t_0+1, \dots, t_1 - 1$ is the solution of equation (6.17) with $\bar{P}_{t_1} = P_1$, it follows that $G_s = \bar{P}_{t^*-s}$ is the solution of (6.18) with $G_{t_0} = P_1$ given. ∎

We also have the following corollary:

Corollary 6.1: Consider the problem of maximizing G_{t_1} subject to $G_{t_0} = P_1$ given and the difference equation (6.18), where the maximization is with respect to $\{H_{s-1}, s = t_0 + 1 \dots, t_1\}$. The maximizing values of H_{s-1} are

$$H_{s-1}^o = [B^T G_{s-1}^o B + Q]^{-1} B^T G_{s-1}^o A$$

where the optimized value G_s^o of G_s obey the "forward" matrix Riccati equation

$$
\begin{aligned}
G_s^o &= A^T G_{s-1}^o A + R + A^T G_{s-1}^o B (B^T G_{s-1}^o) B \\
&\quad + Q^{-1} B^T G_{s-1}^0 A.
\end{aligned} \tag{6.21}
$$

Proof: We leave it as an exercise for the reader to show that theorem 6.4 ad theorem XXX (matrix Ricatti equation) readily imply the corollary. ∎

Now rewrite the difference equation (6.15) Σ_{t+1} as

$$
\begin{aligned}
-\Sigma_{s+1} &= [A^T - C^T K_s^T]^T (-\Sigma_s) [A^T - C^T K_s^T] \\
&\quad + (-V_1) + [K_s^T]^T (-V_2) [K_s^T],
\end{aligned} \tag{6.22}
$$

subject to $\Sigma_{t_0} = \Sigma_0$ given. Evidently maximizing $-\Sigma_{t_1}$ with respect to $\{K_s, s = t_0, t_0 + 1, \ldots, t_1 - 1\}$ is equivalent with minimizing Σ_{t_1}. It immediately follows from corollary 6.1 that the optimal choice of K_s is given by

$$
\begin{aligned}
K_s^T &= [C \Sigma_s^o C^T + V_2]^{-1} C \Sigma_s^o A^T, \\
s &= t_0, t_0 + 1, \ldots, t_1 - 1
\end{aligned} \tag{6.23}
$$

where Σ_s^o is generated from

$$
\begin{aligned}
\Sigma_{s+1}^o = A \Sigma_s^o A^T + V_1 - A \Sigma_s^o C^T& \\
[C \Sigma_s^o C^T + V_2]^{-1} C \Sigma_s^o A^T& \\
s = t_0, t_0 + 1, \ldots, t_1 - 1,&
\end{aligned} \tag{6.24}
$$

with $\Sigma_{t_0}^o = \sum_0$ given.

3. Duality

It is useful at this point to recall the equations that describe the solution of the optimal linear regulator problem:

$$
F_t = (B^T P_{t+1} B + Q)^{-1} B^T P_{t+1} A \tag{6.25}
$$

$$
\begin{aligned}
P_{t-1} &= A^T P_t A + R - A^T P_t B (B^T P_t B + Q)^{-1} B^T P_t A \\
t &= t_0, t_0 + 1, \ldots, t_1 - 1.
\end{aligned} \tag{6.26}
$$

The concept of duality is the key to characterizing the relationship between the two problems and their solutions. Thus, suppose that we have a time invariant linear optimal regulator problem with given matrices, A, B, Q, R, and P_{t_t} and that the parameters t_1 and t_0 are given. Let $\bar{P}_t$ and $\bar{F}_t$ be the solutions for this problem that obey (6.25) and (6.26), for $t = t_0, t_0+1, \ldots, t_1 - 1$.

Now consider creating the optimal linear stochastic observer problem for the system

$$
\begin{aligned}
x^*_{t+1} &= A^T x^*_t + w_{1t+1} \\
y^*_t &= B^T x^*_t + w_{2t}
\end{aligned}
\tag{6.27}
$$

where

$$
E \begin{bmatrix} w_{1t+1} \\ w_{2t} \end{bmatrix} \begin{bmatrix} w_{1t+1} \\ w_{2t} \end{bmatrix}^T = \begin{bmatrix} R & 0 \\ 0 & -Q \end{bmatrix}.
$$

Further, suppose that the optimal linear stochastic observer problem is to be solved for the time period starting from t_0 and ending at $t_1 > t_0$. Let $-\sum_{t_0} = P_1$ be given, where P_1 is the same negative semidefinite matrix used as the terminal value matrix in the optimal linear regulator problem. It follows immediately from equations (6.23), (6.24), on the one hand, and (6.25), (6.26), on the other hand, that the solution to the optimal linear stochastic observer problem is given by

$$
\begin{aligned}
-\Sigma_{t_0} &= \bar{P}_{t_1} = P_1 = -\Sigma_0 \\
-\Sigma_{t_0+1} &= \bar{P}_{t_1-1} \\
-\Sigma_{t_0+2} &= \bar{P}_{t_1-2} \\
&\vdots \\
-\Sigma_t &= \bar{P}_t
\end{aligned}
\tag{6.28}
$$

$$
\begin{aligned}
K^T_{t_0} &= \bar{F}_{t_1-1} \\
K^T_{t_0+1} &= \bar{F}_{t_1-2} \\
&\vdots \\
K^T_{t_1+1} &= \bar{F}_{t_1}.
\end{aligned}
\tag{6.29}
$$

This claim can be established directly by verifying that the solutions (6.28) and (6.29) satisfy (6.23) and (6.24) with the correct boundary condition $\Sigma_{t_0} = -P_1$. Thus, the solution of an optimal linear regulator problem can always be reinterpreted as the solution of a specific optimal linear stochastic observer problem for the *dual* (6.27) of the system for which the regulator problem is solved. These interconnections are usefully summarized in Table 1.

Table 1

Object in Optimal Regulator Problem	**Object in Corresponding Optimal Observer**
A	A^T
B	C^T
R	$-V_1$
Q	$-V_2$
P_{t_0}	$-\Sigma_{t_1}$
P_{t_0+1}	$-\Sigma_{t_1-1}$
.	.
.	.
.	.
P_{t_1-1}	$-\Sigma_{t_0+1}$
P_{t_1}	$-\Sigma_{t_0}$
P_1	$-\Sigma_0$
F_{t_0}	$K_{t_1-1}^T$
F_{t_0+1}	$K_{t_1-2}^T$
.	.
.	.
.	.
F_{t_1-1}	$K_{t_0}^T$
$A - BF_{t_0}$	$A^T - C^T K_{t_1-1}^T$
.	
.	
.	
.	
$A - BF_{t_1-1}$	$A^T - C^T K_{t_0}^T$

4. Convergence and Stability Theorems for the Optimal Observer

For systems that are time invariant, two properties of the limiting behavior of Σ_t and K_t given by (6.23) and (6.24) are desirable. First, it would be desirable if for any initial $\Sigma_0 = \Sigma_{t_0}, \lim_{t\to\infty} \Sigma_t$ exists and is independent of the initial covariance matrix $\sum_0$. Where this property obtains, it follows from (6.23) that $\lim_{t\to\infty} K_t$ exists and is independent of $\sum_0$. Given that $\lim_{t\to\infty} K_t = K$ exists, a second property would be desirable, namely that the matrix $(A - KC)$ be a stable matrix. The steady state observer is given by

(6.30)
$$\hat{x}_{t+1} = (A - KC)\hat{x}_t + Ky_t$$
or
$$\hat{x}_{t+1} = A\hat{x}_t + Ky_t - C\hat{x}_t).$$

Notice that the system (6.30) has the solution

(6.31)
$$\hat{x}_{t_0+j} = (A - KC)^j \hat{x}_{t_0} + \sum_{i=1}^{j} (A - KC)^{j-i} K y_{t_0+i-1}.$$

If the eigenvalues of $(A - KC)$ are bounded in modulus by unit, (6.31) expresses $\hat{x}_{t_0+j}$ as a matrix distributed lag of $y_{t_0+j-1}, \ldots, y_{t_0}$ with an initial condition whose effect approaches zero as $j \to \infty$. The steady states observer is said to be *asymptotically stable* if $(A - KC)$ is a stable matrix.

The fact that the stochastic linear optimal observer problem is dual to the linear optimal regulator problem, means that we can simply reinterpret the sufficient conditions for P_{t_0} to converge as $t_0 \to \infty$ in order to deduce sufficient conditions for $\sum_{t_1}$ to converge as $t_1 \to \infty$. Similarly, from the conditions on the linear regulator problem sufficient for the steady state closed loop system matrix $(A - BF)$ to be stable, we can immediately deduce conditions on the observer problem sufficient for $(A - KC)$ to be stable.

We proceed to state several theorems for the stochastic linear optimal observer problem that follow by duality from corresponding theorems for the optimal linear regulator problem. Corresponding to theorem — we have:

Theorem 6.5: Consider the stochastic optimal linear observer problem with $\Sigma_0 = 0$. Assume that the pair (A, C) is reconstructible. Then the reconstruction error covariance matrix Σ_{t_1} calculated from the Riccati equation (6.24) converges as $t_1 \to \infty$.

Proof: The pair (A, C) is reconstructible if and only if the rank of the reconstructibility matrix

$$Q = \begin{bmatrix} C \\ CA \\ \cdot \\ \cdot \\ \cdot \\ CA^{n-1} \end{bmatrix}$$

is n. Consider the linear optimal regulator problem corresponding to the optimal observer problem (see Table 1). In the corresponding linear optimal regular problem, the controllability matrix is $[C^T, A^T, C^T, \ldots, A^T, C^{T\,n-1}]$, which has rank n. Thus the pair $(\tilde{A}, \tilde{C}) = (A^T, C^T)$ is completely controllable, implying by virtue of theorem — that starting from $P_{t_1} = 0, \lim_{t_0 \to -\infty} P_{t_0}$ exists. Therefore, by virtue of duality $\lim_{t_1 \to \infty} \Sigma(t_1) = \lim_{t_0 \to -\infty} P_{t_0}$ exists. ∎

Corresponding to theorem — we have

Theorem 6.6: Consider the stochastic optimal linear observer problem with $\sum_{t_0} = 0$. Let the $(n \times n)$ positive semidefinite matrix V_1 be expressed as $G^T G$ where G is $(r \times n)$, $r \leq n$ and r is the rank of V_1. Assume that the pair (A, C) is *reconstructible*, and that the pair (A, G) is *controllable*. Then the steady state matrix $(A - KC)$ is stable.

Proof: It is readily verified that (A, C) is reconstructible if and only if (A^T, C^T) is controllable, and that (A, G) is controllable if and only if (A^T, G^T) is reconstructible. Set $\tilde{A} = A^T, \tilde{B} = C^T, \tilde{G} = G^T, -\tilde{R} = V_1,$ and $-\tilde{Q} = V_2$. Consider the optimal regulator problem for the system $\tilde{G}, \tilde{B}, \tilde{G}, \tilde{R},$ and $\tilde{Q}$ with $P(t_1) = 0$. From theorem it follows that the steady state closed loop matrix $(\tilde{A} - \tilde{B}\tilde{F})$ is stable. By virtue of duality, we have that $\tilde{K}^T = \tilde{F}$ and $(\tilde{A} - \tilde{B}\tilde{F}) = (A^T - C^T K^T)$. This implies that $(A - KC)$ is a stable matrix. ∎

By utilizing theorem (), the hypotheses of theorem 3.6 can be weakened from assuming that the pair (A, G) is controllable to assuming that the pair (A, G) is stabilizable. Under this assumption it remains true that the steady state matrix $(A - KC)$ is stable.

We also have the following counterpart to theorem — :

Theorem 6.7: Consider the stochastic linear optimal observer problem. Assume that sufficient conditions are satisfied so that iterations on the Riccati equation (6.15) starting from Σ_{t_0} = converge, and that the associated steady state matrix $(A - KC)$ is stable. Then for any positive semidefinite initial covariance matrix Σ_{t_0}, iterations on the matrix Riccati equation (6.24) converge to the same positive semidefinite matrix Σ, i.e., the limit point described in theorem 6.5.

Proof: Exercise.

By this time the reader will have understood that by virtue of duality, all of the theorems stated for the optimal linear regulator problem have interesting counterparts for the optimal linear stochastic observer problem. We invite the reader to state and prove the counterparts to theorems — — —.

5. An Example: (Muth [], Friedman [], Cagan [])

An agent is interested in making inferences about a random variable θ_t which obeys the first-order autoregressive process

$$\theta_{t+1} = \rho\theta_t + \varepsilon_{t+1},$$

where ε_t is a white noise with $E\varepsilon_t = 0, E\varepsilon_t^2 = \sigma_\varepsilon^2$ for all t. The agent observes at time t the record of noise corupted signals $z_t, z_{t-1}, \ldots, z_{t_0}$, where

$$z_t = \theta_t + u_t$$

and where u_t is a serially uncorrelated random process with $Eu_t = 0, Eu_t^2 = \sigma_u^2$. We also assume that $Eu_t\varepsilon_s = 0$ for all t and s. The agent desires to estimate θ_{t+1} on the basis of information he possesses at t. At time t_0, θ_{t_0} is (believed to be) distributed with mean $\bar{\theta}_0$ and variance Σ_0.

This problem fits into the stochastic linear optimal observer problem with the following identifications. We set

$$x_t = \theta_t$$

$$w_{1t+1} = \varepsilon_{t+1}, y_t = z_t$$

$$w_{2t} = u_t$$

$$V = E\begin{bmatrix} w_{1t+1} \\ w_{2t} \end{bmatrix}\begin{bmatrix} w_{1t+1} \\ w_{2t+1} \end{bmatrix}^T = \begin{bmatrix} \sigma_\varepsilon^2 & 0 \\ 0 & \sigma_u^2 \end{bmatrix} = \begin{bmatrix} V_1 & 0 \\ 0 & V_2. \end{bmatrix}$$

$$A = \rho$$

$$C = 1$$

$$\hat{\theta}_t = \hat{E}(\theta_t \mid z_{t-1}, z_{t-2}, \ldots, z_{t_0})$$

$$\Sigma_t = E(\theta_t - \hat{\theta}_t)$$

The recursive equations defining the filter are

$$\begin{aligned}\Sigma_{t+1} &= \rho^2 \Sigma_t + \sigma_e^2 \\ &\quad - \rho^2 \Sigma_t^2 [\Sigma_t + \sigma_u^2]^{-1}, \quad \text{with} \quad \Sigma_{t_0} = \Sigma_0.\end{aligned} \tag{6.32}$$

$$K_t = \rho \Sigma_t \left[\Sigma_t + \sigma_u^2\right]^{-1}. \tag{6.33}$$

The optimal observer is

$$\hat{\theta}_{t+1} = (\rho - K_t)\hat{\theta}_t + K_t z_t. \tag{6.34}$$

We can readily verify that the pair $(A, C) = (\rho, 1)$ is *reconstructible*, and that the pair $(A, G) = (\rho, \sigma_\epsilon)$ (where $GG^T = V_1$) is *controllable*. Therefore from theorems 6.5–6.7, we know that $\lim_{t\to\infty} \Sigma_t$ exists, that $\lim_{t\to\infty} K_t = K$ exists, and that the steady state matrix $(A - KC)$ is stable.

The steady state observer is

$$\hat{\theta}_{t+1} = (\rho - K)\hat{\theta}_t + K z_t$$

or

$$\hat{\theta}_{t_0+j} = (1 - K)^j \hat{\theta}_{t_0} + K \sum_{i=1}^{j} (1 - K)^{j-i} z_{t_0+i-1}.$$

where recall that $\hat{\theta}_{t_0+j} = E\theta_{t_0+j} \mid [z_s, s = t_0, \ldots, t_0 + j - 1]$.

6. The Optimal Linear Regulator Problem with Hidden State Variables

We consider the problem of maximizing the criterion

$$E\left[\sum_{t=t_0}^{t_1-1} \{x_t^T R x_t + u_t^T Q u_t\} + x_{t_1}^T P_{t_1} x_{t_1}\right] \tag{6.32}$$

subject to the law of motion

$$x_{t+1} = A x_t + B u_t + w_{1t+1}$$

where w_{1t+1} is a vector white noise with $Ew_{1t}w_{1t}^T = V_1$. The state vector x_t is not observed by the problem-solver. Instead, at t the problem-solver sees $\{y_s, u_s; s \leq t\}$ where

$$y_t = Cx_t + w_{2t}$$

where w_{2t} is a vector white noise with covariance matrix $Ew_{2t}w_{2t}^T = V_2$. We also assume

$$Ew_{1t+1}w_{2s}^T = \begin{cases} V_{12}, & t = s \\ 0 & t \neq s \end{cases}.$$

The criterion (6.32) is to be maximized over feedback laws making u_t a function of $(y_s, u_{s-1}; s \leq t)$.

We shall show that the solution of this problem can be obtained in two steps. First, solve the standard optimal linear regulator problem that results from assuming that x_t itself is observed, obtaining the sequence of linear feedback rules $u_t = -F_t x_t$. Second, from the linear-least-squares estimator $\hat{x}_t$ of the hidden state x_t using the Kalman-filter

$$\hat{x}_{t+1} = (A - K_t C)\hat{x}_t + K_t y_t$$

Then the optimal solution for the problem (6.32) is to use the control law

$$u_t = -F_t \hat{x}_t \tag{6.33}$$

i.e., to feedback on the optimally reconstructed state as though it were the actual state. This structure of the solution indicates the sense in which the optimization (linear regulator) problem and the state reconstruction (Kalman filtering) problem can be solved separately in solving the general linear regulator problem (6.32) with hidden state variables. This structure of the solution is said to mean that it satisfies a *separation principle*.

To prove the separation principle property, we begin by noting that for the optimally reconstructed state $\hat{x}_t$

$$\begin{aligned} Ex_t^T R x_t &= E\,[x_t - \hat{x}_t + \hat{x}_t]^T R [x_t - \hat{x}_t + \hat{x}_t] \\ &= E\,\{[x_t - \hat{x}_t]^T R [x_t - \hat{x}_t]\} \\ &\quad + 2E\,\{[x_t - \hat{x}_t]^T R \hat{x}_t\} + E\hat{x}_t^T R \hat{x}_t. \end{aligned} \tag{6.34}$$

Letting $\Sigma_t = E(x_t - \hat{x}_t)(x_t - \hat{x}_t)^T$, recall that

$$E(x_t - \hat{x}_t^T)R(x_t - \hat{x}_t) = trR\Sigma_t.$$

Further, by the orthogonality principle, we have

$$E\ (x_t - \hat{x}_t)^T R\hat{x}_t = tr\left[\hat{x}_t(x_t - \hat{x}_t)^T R\right] = 0\ .$$

Therefore, we have

$$E\ x_t^T R x_t = tr R\Sigma_t + E\hat{x}_t^T R\hat{x}_t. \tag{6.35}$$

Using (6.35) and the analogous expression for $E\ x_t^T P_{t_1} x_{t_1}$ in (6.32) gives the criterion function

$$E\left[\sum_{t=t_0}^{t_1-1}\{\hat{x}_t^T R\hat{x}_t + u_t^T Q u_t\} + \hat{x}_{t_1}^T P_{t_1}\hat{x}_{t_1} + tr\Big\{\sum_{t=t_0}^{t_1-1} R\Sigma_t + P_{t_1} E\varepsilon_{t_1}\Big\}\right] \tag{6.36}$$

The last term in braces is independent of the controls u_t, since the problem solver is assumed to see current and lagged controls, so they don't confound his reconstruction problem. The last terms in braces evidently depends only on the statistics of the optimal reconstruction problem, and furthermore is maximized by the optimal observer, since the Σ_t sequence is minimized. Our problem is now to maximize (6.36) subject to the following law of motion for the reconstructed state:

$$\hat{x}_{t+1} = (A\hat{x}_t + Bu_t) + K_t\left[y_t - C\hat{x}_t\right] \tag{6.37}$$

It was established above that $(y_t - C\hat{x}_t)$ is a vector white noise, so that maximizing (6.36) subject to (6.37) is a standard stochastic linear optimal regulator problem with state vector $\hat{x}_t$ known, with system matrices (A, B, R, Q), and with noise statistics given by

$$K_t E(y_t - C\hat{x}_t)\ (y_t - C\hat{x}_t)^T K_t^T$$

The optimal solution of this problem is of the form

$$u_t = -F_t\hat{x}_t.$$

This concludes the proof that our problem possesses the separation principle property.

Next we study the behavior of the system under conditions in which both F_t and K_t converge to limiting values. The asymptotic closed loop system governing the $(2n \times 1)$ system of variables $(x_t, \hat{x}_t)$ is then

$$\begin{aligned} x_{t+1} &= Ax_t - BF\hat{x}_t + w_{1t+1} \\ \hat{x}_{t+1} &= (A - KC)\hat{x}_t - BF\hat{x}_t + KC(x_t) + Kw_{2t}. \end{aligned} \tag{6.38}$$

It is useful to express this system in terms of the variables $x_t, x_t - \hat{x}_t$. (This can be accomplished by subtracting the second equation from the first.) In terms of these variables the system is

$$(6.39) \qquad \begin{bmatrix} x_{t+1} \\ x_{t+1} - \hat{x}_{t+1} \end{bmatrix} = \begin{bmatrix} A - BF & BF \\ 0 & A - KC \end{bmatrix} \begin{bmatrix} x_t \\ x_t - \hat{x}_t \end{bmatrix} + \begin{bmatrix} w_{1t+1} \\ w_{1t+1} - Kw_{2t} \end{bmatrix}$$

Since the system matrix of (6.38) is related to that of (6.39) by a similarity transformation, it shares common eigenvalues with that of (6.39). From the block triangular structure of the system matrix in (6.39), it follows that its eigenvalues are the eigenvalues of $A - BF$ and those of $A - KC$. This property is known as the *eigenvalue separation theorem.*

Theorem 6.8 (Eigenvalue Separation): The variables $\begin{bmatrix} x_t \\ \hat{x}_t \end{bmatrix}$ are governed by a linear system with a transition matrix whose characteristic values are those of $(A - BF)$ and $(A - KC)$ jointly.

7. Econometric Estimation

We now consider the problem of estimating the free parameters of a model of the form (6.38), namely,

$$(6.40) \qquad \begin{bmatrix} x_{t+1} \\ \hat{x}_{t+1} \end{bmatrix} = \begin{bmatrix} A & -BF \\ KC & A - KC - BF \end{bmatrix} \begin{bmatrix} x_t \\ \hat{x}_t \end{bmatrix} + \begin{bmatrix} w_{1t+1} \\ Kw_{2t} \end{bmatrix}$$

where

$$V \equiv E \begin{bmatrix} w_{1t+1} \\ Kw_{2t} \end{bmatrix} \begin{bmatrix} w_{1t+1} \\ Kw_{2t} \end{bmatrix}^T = \begin{bmatrix} V_1 & V_3 K^T \\ K V_3^T & K V_2 K^T \end{bmatrix}$$

The model is subject to the extensive cross-equation restrictions

$$(6.41) \qquad F = (B^T P B + Q)^{-1} B^T P A$$

$$(6.42) \qquad K = (A \Sigma C^T (C \Sigma C^T + V_2)^{-1}$$

where P is the unique negative semi-definite solution of

$$(6.43) \qquad P = (A^T P A + R - A^T P B (B^T P B + Q)^{-1} B^T P A$$

and Σ is the unique positive semi-definite solution of

$$\Sigma = A\Sigma A^T + V_1 - A\Sigma C^T(C\Sigma C^T + V_2)^{-1}C\Sigma A^T \tag{6.44}$$

Equation (6.40) is a vector first-order linear difference equation in the variables $(x_t, \hat{x}_t)$, some subset of which we assume that the econometrician observes. The econometrician's problem is to estimate the free parameters of agents' objective functions and constraints. From the econometrician's viewpoint, the free parameters of the model are the free parameters in $\theta \equiv (R, Q, A, B, C, V_1, V_2, V_3)$. The parameters of F and K enter the "closed loop" law of motion (6.40), but are not free parameters, instead being functions of the deep free parameters in the list θ. Thus, the model (6.40) to be estimated is linear in the variables but highly nonlinear in the deep parameters of agents' objective functions and constraints. These nonlinear restrictions are characterized by equations (6.41), (6.42), (6.43) and (6.44).

The general theory of estimation can be stated compactly and simply. The model formed by (6.40)–(6.44) determines the second moments of the joint $(x_t \hat{x}_t)$ process as functions of the free parameters. The idea behind all alternative estimators is to choose the free parameters in θ so that the sample moments of the data on which the econometrician has observations fits the theoretical moments implied by the model as closely as possible. Alternative estimators differ in implicitly choosing different measures of fit. We turn briefly to the maximum likelihood estimator, which is straightforward to describe.

First, we rewrite equation (6.40) as

$$y_{t+1} = \tilde{A} y_t + \varepsilon_{t+1} \tag{6.45}$$

where

$$y_t = \begin{bmatrix} x_t \\ \hat{x}_t \end{bmatrix}, \varepsilon_{t+1} = \begin{bmatrix} w_{1t+1} \\ K w_{2t} \end{bmatrix}$$

$$\tilde{A} = \begin{bmatrix} A & -BF \\ KC & A - KC - BF, \end{bmatrix}$$

We assume that sufficient conditions are met that the eigenvalues of $\tilde{A}$, which equal those of $A - BF$ and $A - KC$, are less than unity in modulus. We define the matrix covariogram of the y_t process as the sequence of $(2n \times 2n)$ matrices

$$R_y(\tau) = E y_t y_{t-\tau}^T, \quad \text{integer} \quad \tau.$$

The z-transform of the covariogram is defined as

$$(6.46) \qquad S_y(z) = \sum_{\tau=-\infty}^{\infty} R_y(\tau) z^{\tau}$$

where $R(\tau)$ is recoverable from $S_y(z)$ by the inversion formula

$$R_y(\tau) = \frac{1}{2\pi i} \int_{\Gamma} S_y(z) z^{-\tau} \frac{dz}{z}$$

where the integral is a contour integral and Γ denotes the unit circle. For a model of the form (6.45), the $z-$ transform of the autocovariogram can be shown to be

$$S_y(z) = (I - \tilde{A}z)^{-1} V (I - \tilde{A}^T z^{-1})^{-1}$$

Let the eigenvalues of $\tilde{A}$ be $\lambda_1, \ldots, \lambda_r$ where $r = 2n$. We have assumed that $\mid \lambda_j \mid < 1$ for all j, and assume also that the λ_j's are distinct. Then by using a matrix partial fractions representation of $S_y(z)$, it can be shown that

$$(6.47) \qquad S_y(z) = \sum_{j=1}^{r} \frac{W_j}{1 - \lambda_j z} + \sum_{j=1}^{r} \frac{W_j^T \lambda_j z^{-1}}{1 - \lambda_j z^{-1}}$$

Expressing $S_y(z)$ as

$$S_y(z) = \frac{1}{det(I - \tilde{A}z) det(I - \tilde{A}^T z^{-1})} \, adj(I - \tilde{A}z) V adj(I - \tilde{A}^T z^{-1})$$

We note that $S_y(z)$ has poles at the zeroes of $det(I - \tilde{A}z)$, that the zeroes of $det(I - \tilde{A}^T z^T)$ are the eigenvalues of $\tilde{A}$, and that the zeroes of $det(I - \tilde{A}z)$ are the reciprocals of the eigenvalues of $\tilde{A}$. Writing $det(I - \tilde{A}z) = \lambda_0 (1 - \lambda_1 z), \ldots (1 - \lambda_r z)$ we have

$$(6.48) \qquad S_y(z) = \frac{1}{\lambda_0^2 \Pi_{j=1}^{r} (1 - \lambda_j z) \Pi_{k=1}^{r} (1 - \lambda_k z^{-1})} \; adj(I - \tilde{A}z) V \, adj(I - \tilde{A}^T z^{-1})$$

Now seek a matrix partial fractions representation of the form

$$(6.49) \qquad S_y(z) = \sum_{j=1}^{r} \frac{W_j}{1 - \lambda_j z} + \frac{V_j z^{-1}}{1 - \lambda_j z^{-1}},$$

Equating (6.48) and (6.49) and multiplying both sides by $\lambda_0^2 \Pi_{j=1}^{r} (1 - \lambda_j z) \Pi_{k=1}^{r} (1 - \lambda_k z^{-1})$, then taking limits as $z \to \lambda_j^{-1}$ and $z \to \lambda_j$, respectively, gives

$$W_j = [\lambda_0^2 \Pi_{h=j} (1 - \lambda_h \lambda_j^{-1}) \Pi_{k=1}^{r} (1 - \lambda_k \lambda_j)]^{-1}) \; adj \; \{I - \tilde{A}_j^{-1}\} \; V adj \; \{I - A^T \lambda_j\}$$

$$V_j = \lambda_j W_j^T$$

Substituting these formulas into (6.49) gives formulas equivalent with (6.44) and () of the text. Where

(6.50) $$W_j = \lim_{z \to \lambda_j^{-1}} (1 - \lambda_j z)[(I - \tilde{A}z)^{-1} V (I - \tilde{A}^T z^{-1})^{-1}]$$

From (6.47), the covariogram can be immediately obtained as

(6.51) $$R_y(\tau) = \begin{bmatrix} \sum_{j=1}^{\tau} W_j \lambda_j^{\tau}, & \tau \geq 0 \\ \sum_{j=1}^{\tau} W_j^T \lambda_j^{|\tau|}, & \tau \leq 0 \end{bmatrix}$$

Equation (6.47), (6.50), and (6.51) give the theoretical second moments of the vector process $y_t = (x_t^T, \dot{x}_t^T)$ as a function of the free parameters that underlie $\tilde{A}$ and V.

Now suppose that the econometrician has data on some subset of p variables $\bar{y}_t = Dy_t$ where $p < 2n$, and where D is a $(p \times 2n)$ matrix. Then it is readily verified that the matrix covariogram of $\bar{y}_t$, call it $R_{\bar{y}}(\tau) = E\bar{y}_t\bar{y}_{t-\tau}^T$, is given by

$$R_{\bar{y}}(\tau) = DR_y(\tau)D^T$$

(6.52) $$R_y(\tau) = \begin{bmatrix} \Sigma_{j=1}^{\tau} DW_j D^T \lambda_j^{\tau}, & \tau \geq 0 \\ \Sigma_{j=1}^{\tau} DW_j^T D^T \lambda_j^{|\tau|}, & \tau \leq 0. \end{bmatrix}$$

Equation (6.52) gives the covariogram of the variables on which the econometrician has data as a function of D and the W_j's, λ_j's, which in turn are functions of the deep parameters of the model.

Now define the stacked vector of observations

$$\bar{y}_t = \begin{bmatrix} \bar{y}_1 \\ \bar{y}_2 \cdot \\ \cdot \\ \cdot \\ \bar{y}_T \end{bmatrix}$$

Define the theoretical covariance matrix of $\bar{y}_T$,

$$\Gamma_T(\theta) = E\bar{y}_T\bar{y}_T^T,$$

whose elements are components of $R\bar{y}(\tau)$ and can be filled in as functions of the deep parameters θ of the model by using (6.52). Then the normal likelihood function of the sample $\bar{y}_T$ is given by

(6.53) $$L_T = -\frac{1}{2}Tp\ log\, 2\pi - \frac{1}{2}log\ det\, \Gamma_T(\theta) - \frac{1}{2}y_T^{-T}\Gamma_T^{-1}(\theta)\bar{y}_T$$

Estimation proceeds by choosing the free parameters of θ to maximize (6.52) subject to the cross-equation restrictions given by (6.40), (6.41), (6.42), (6.43), (6.44), (6.45), (), (6.50), and (6.52).

Exercises

1. The "true" money supply follows the stochastic process

$$M_t = \lambda M_{t-1} + u_t$$

where $Eu_t = 0, u_t = [M_t - EM_t \mid M_{t-1}, \ldots]$, u_t has finite variance , and u_t is serially uncorrelated. But "true" money is reported only with a two-period lag; what is reported immediately is a preliminary estimate of money m_t , governed by

$$m_t = M_t + \varepsilon_t,$$

where ε_t has zero mean, is serially uncorrelated, has finite variance, and $Eu_t\varepsilon_s = 0$ for all t and s. Suppose that the system has been operating for a long time, so that it is a good approximation to assume that λ and the variances of ε_t and u_t are known.

(*a*) Show how to compute the linear least squares estimators of m_{t+1} and m_t given information know at time t; i.e., compute $E_t m_{t+1}$ and $E_t m_t$ where E is the linear least squares projection operator. Hint: use the Kalman filter or full order observer algorithm, and define x_t, y_t, A, B, C.

(*b*) Is V_2 positive definite? If not, does this create problems with the algorithm you outlined? Can you think of a way of coaxing the full order observer algorithm of class to give a good approximate answer? (Hint: think of a way of approximating the true V_2 by a positive definite V_2). What is the interpretation of your approximate solution?

(*c*) With V_2 positive definite, prove that for the approximate system

(*i*) Iterations on the matrix Riccati difference equation for $\sum(t)$ converge as $t \to \infty$.

(*ii*) The steady state matrix $(A - KC)$ is stable, regardless of the value of λ.

2. Consider the state space system

$$x_{t+1} = A_t x_t + B_t u_t + w_{1t+1}$$

$$y_t = C_t x_t + w_{2t}$$

as described in the text. Describe carefully how the following examples fit into the state space framework (i.e., for each example, you must define x_t, A_t, B_t, u_t, w_{1t+1}, y_t, C_t, and w_{2t}).

(a) An autoregressive process

$$z_t = a_1 z_{t-1} + a_2 z_{t-2} + \cdots + a_n z_{t-n} + \varepsilon_t$$

where ε_t is fundamental for $\{z_t\}$, and the roots of $(1 - a_1 z - a_2 z^2 - \ldots - a_n z^n) = 0$ are outside the unit circle.

(b) A moving average process

$$z_t = c_0 \varepsilon_t + c_1 \varepsilon_{t-1} + \cdots + c_n \varepsilon_{t-n}$$

where ε_t is a white noise that is fundamental for z_t. (Hint: define the *state* vector as $x_t = (\varepsilon_t, \varepsilon_{t-1}, \cdots, \varepsilon_{t-1}))$.

(c) A mixed moving average autoregressive process

$$z_t = a_1 z_{t-1} + \varepsilon_t + b_1 \varepsilon_{t-1}$$

where $| a_1 |< 1, | b_1 |< 1$, and ε_t is fundamental for z_t.

(d) A regression model

$$Y_t = X_t \beta + \varepsilon_t \qquad t = 1, \ldots, T$$

where ε_t is a white noise, X_t are fixed regressors, and β is a vector of regression coefficients.

(e) A moving coefficients regression model

$$Y_t = X_t \beta_t + \varepsilon_t$$

where $\beta_t = \beta_{t=1} + u_t$, where u_t is a white noise and the other symbols are as defined in (d).

Chapter 7

Linear Dynamic Equilibrium Models

1. Introduction

This chapter discusses two alternative ways of solving linear dynamic equilibrium models. The first, which we dub the Kydland-Prescott method, is a recursive method that works even in the presence of dynamic externalities and other distortions. The second method, which is Lucas and Prescott's, exploits the equivalence between equilibrium and optimality. This second method will not work for environments with some distortions that can be handled by the Kydland-Prescott method.

We describe these mthods in the context of a concrete model, namely, a version of Lucas and Prescott's model of investment under uncertainty with adjustment costs. We go on to apply these methods to a two sector model of "corn-hog" equilibrium dynamics.

2. The Kydland-Prescott Method

This is a model of an industry in which n identical competitive firms employ a single productive input, capital, to produce a single output. The industry demand curve for output at time t is

$$p_t = A_0 - A_1 Y_t + u_t, \qquad A_0 > 0, A_1 > 0 \tag{7.1}$$

where p_t is output price at t, Y_t is industry output, and u_t is a shock to demand. The output of each firm is $y_t = f_0 k_t$ where k_t is the firm's capital stock at time t and $k_0 > 1$. The industry-wide capital stock is $K_t = nk_t$, and the industry-wide output is $Y_t = ny_t = nf_0 k_t = f_0 K_t$. The firm pays a rental w_t per unit of capital at time t. The rental process w_t is assumed to follow the law

$$w_t = \lambda_0 + \lambda_1 w_{t-1} + \lambda_2 K_{t-1} + a_{wt} \tag{7.2}$$

where a_{wt} is a serially uncorrelated random process with mean zero. It is important to note that with $\lambda_2 \neq 0$, (7.2) permits feedback or Granger causality from the market-wide capital stock K to the rentals process w. At time t, the market wide capital stock K_t is assumed to follow the linear law of motion.

$$K_{t+1} = h_0^t + h_1^t K_t + h_2^t w_t + h_3^t u_t, \tag{7.3}$$

where notice that the coefficients in (7.3) are permitted to depend on time. We suppose that the demand shock u_t follows the autoregressive process

$$u_{t+1} = \alpha u_t + a_{ut+1} \tag{7.4}$$

where a_{ut+1} is a serially uncorrelated random process with mean zero that is fundamental for u.

The individual firm is supposed to maximize

$$E_{t_0} \sum_{t=t_0}^{t_1} \beta^{(t-t_0)} \left\{ p_t f_0 k_t - w_t k_t - \frac{d}{2}\left(k_{t+1} - k_t\right)^2 \right\} \tag{7.5}$$
$$d > 0, \qquad 0 < \beta < 1$$

subject to k_{t_0} given, and subject to knowledge of the laws of motion (7.2), (7.3), and (7.4), and the demand curve (7.1). In (7.5), $\frac{d}{2}(k_{t+1} - k_t)^2$ represents costs of adjusting the capital stock rapidly, while β is a discount factor. At time t, the firm is supposed to choose k_{t+1} as a function of the state variables it knows, namely $\{k_t, K_t, \omega_t, u_t\}$. In (7.4) $E_{t_0}(\cdot) = E(\cdot \mid k_{t_0}, K_{t_0}, \omega_{t_0}, u_{t_0})$, where E is the mathematical expectations operator. The solution of this problem will be a sequence of linear contingency plans

$$k_{t+1} = d_0^t + d_1^t k_t + d_2^t w_t + d_3^t u_t + d_4^t K_t, \tag{7.5}$$
$$t = t_0, t_{0+1}, \ldots, t_1 ,$$

where the coefficients d_j^t are in general dependent on time.

Equilibrium requires that the choice (7.6) of the representative firm imply the aggregate law of motion (7.3) assumed by firms in maximizing (7.5). Multiplying both sides of (7.6) by n gives

$$K_{t+1} = nd_0^t + (d_1^t + nd_4^t)K_t + (nd_2^t)w_t + (nd_3^t)u_t. \tag{7.7}$$

Since (7.7) must be identically equal to (7.3), we have the equilibrium conditions

$$\begin{aligned} h_0^t &= nd_0^t \\ h_1^t &= d_1^t + nd_4^t \\ h_2^t &= nd_2^t \\ h_3^t &= nd_3^t \end{aligned} \tag{7.8}$$

Formally we can define a rational expectations equilibrium as follows.

Definition 7.1: *A rational expectations equilibrium* is a pair of sequences $\bar{h}(t_0, t_1) = \{h_0^t, h_1^t, h_2^t, h_3^t;\ t = t_0, \ldots, t_1\}$ and $\bar{d}(t_0, t_0) = \{d_0^t, d_1^t, d_2^t, d_3^t, d_4^t;\ t = t_0, \ldots, t_1\}$ such that

(*a*) Given $\bar{h}(t_0, t_1)$ as the law of motion in (7.3), $\bar{d}(t_0, t_1)$ in (7.6) maximizes the representative firm's expected present value (7.5).

(*b*) Market clearing and the firm's choice of $\bar{d}(t_0, t_1)$ imply that $\bar{h}(t_0, t_1)$ gives the aggregate law of motion, i.e., equations (7.8) hold.

Let us substitute $[A_0 - A_1 f_0 K_t + u_t]$ for p_t in the objective function (7.5). Then the firm's problem is equivalent with finding a sequence of value functions $V^t(k_t, w_t, u_t, K_t, 1)$ that satisfy Bellman's functional equation

$$
\begin{aligned}
V^t(k_t, w_t, u_t, K_t, 1) = \max_{k_{t+1}} \Big\{ & [A_0 - A_1 f_0 K_t + u_t] f_0 k_t - w_t k_t \\
& - \frac{d}{2}(k_{t+1} - k_t)^2 + \beta E_t V^{t+1}(k_{t+1}, w_{t+1}, u_{t+1}, K_{t+1},\ 1) \Big\}
\end{aligned} \tag{7.9}
$$

where the maximization is subject to the given laws of motion

$$
\begin{aligned}
K_{t+1} &= h_0^t + h_1^t K_t + h_2^t w_t + h_3^t u_t \\
w_{t+1} &= \lambda_0 + \lambda_1 w_t + \lambda_2 K_t + a_{wt+1} \\
u_{t+1} &= \alpha u_t + a_{ut+1}
\end{aligned} \tag{7.10}
$$

In (7.9) it is assumed that $v^{t_1+1}(k_{t_1+1}, w_{t_1+1}, u_{t_1+1}, K_{t_1+1}, 1) \equiv 0$. This is a linear regulator problem that can be solved by standard methods. Define the state vector $X_{t+1} = [k_t, w_t, u_t, K_t, 1]'$, and the control $v_t = k_{t+1} - k_t$. Then the law of motion is

$$
\begin{bmatrix} k_{t+1} \\ w_{t+1} \\ u_{t+1} \\ K_{t+1} \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & \lambda_2 & \lambda_0 \\ 0 & 0 & \alpha & 0 & 0 \\ 0 & h_2^t & h_3^t & h_1^t & h_0^t \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} k_t \\ w_t \\ u_t \\ K_t \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} v_t + \begin{bmatrix} 0 \\ a_{wt+1} \\ a_{ut-1} \\ 0 \\ 0 \end{bmatrix}
$$

or

$$
X_{t+1} = A_t X_t + B v_t + a_{t+1}.
$$

Define the quadratic form $X_t^T R X_t$ as

$$
\begin{bmatrix} k_t \\ w_t \\ u_t \\ K_t \\ 1 \end{bmatrix}^T \begin{bmatrix} 0 & -\frac{1}{2} & +\frac{f_0}{2} & -\frac{A_1 f_0^2}{2} & \frac{f_0 A_0}{2} \\ -\frac{1}{2} & 0 & 0 & 0 & 0 \\ \frac{f_0}{2} & 0 & 0 & 0 & 0 \\ -\frac{A_1 f_0^2}{2} & 0 & 0 & 0 & 0 \\ \frac{f_0 A_0}{2} & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} k_t \\ w_t \\ u_t \\ K_t \\ 1 \end{bmatrix}
$$

Define $Q = -\frac{d}{2}$. Then the firm's problem is the discounted optimal linear regulator problem, to maximize

$$E_{t_0} \sum_{t=t_0}^{t_1} \beta^{(t-t_0)} \left(x_t^T R x_t + v_t^T Q v_t \right),$$

subject to

$$x_{t+1} = A_t x_t + B v_t + a_{t+1}$$

The maximization is over linear contingency plans of the form

$$v_t = -F_t x_t, \qquad t = t_0,\, t_0 + 1, \ldots, t_1.$$

Since $v_t = k_{t+1} - k_t$, we have that

$$F_t = [1 - d_1^t, -d_2^t, -d_3^t, -d_4^t, -d_0^t] \tag{7.11}$$

For this problem with A_t matrices taken as given by the firm, the solution is given by

$$F_t = \beta(\beta B' P_{t+1} B + Q)^{-1} B' P_{t+1} A_t \tag{7.12}$$

where $\{P_t\}$ is computed from the matrix Riccati difference equation,

$$P_t = \beta A_t' P_{t+1} A_t + R_t - \beta^2 A_t' P_t B [\beta B' P_{t+1} B + Q]^{-1} B' P_{t+1} A_t, \tag{7.13}$$

starting from the terminal condition $P_{t_1+1} = 0$. It is revealing to use (7.11) in (7.12) and to write out A_t explicitly to get

$$-[(1 - d_1^t), -d_2^t, -d_3^t, -d_4^t, -d_0^t] =$$

$$\beta(\beta B' P_{t+1} B + Q)^{-1} B' P_{t+1} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & \lambda_2 & \lambda_0 \\ 0 & 0 & \alpha & 0 & 0 \\ 0 & h_2^t & h_3^t & h_{11}^t & h_0^t \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{7.14}$$

Recalling (7.8) we have

$$\begin{aligned} h_0^t &= n d_0^t \\ h_1^t &= d_1^t + n d_4^t \\ h_2^t &= n d_2^t \\ h_3^t &= n d_3^t \end{aligned} \tag{7.8}$$

Given P_{t+1}, equations (7.14) and (7.8) are nine linear equations in the nine variables d_0^t, d_1^t, d_2^t, d_3^t, d_4^t, h_0^t, h_1^t, h_2^t, h_3^t.

Equations (7.13), (7.14), and (7.8) provide the recursive algorithm used by Kydland and Prescott. At t_1, the nine linear equations (7.14) and (7.8) are solved jointly for the nine variables $\{h_0^t, h_1^{t_1}, h_2^{t_1}, h_3^{t_1}, d_0^{t_1}, d_1^{t_1}, d_2^{t_1}, d_3^{t_1}, d_4^{t_1}\}$, with $P_{t_1+1} = 0$. The solution for the $h_j^{t_1}$'s determines the matrix A_{t_1}, where recall that

$$A_t = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & \lambda_2 & \lambda_0 \\ 0 & 0 & \alpha & 0 & 0 \\ 0 & h_2^t & h_3^t & h_1^t & h_0^t \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Given A_{t_1}, P_{t_1} is calculated from the Riccati difference equation (7.13). Then (7.14) and (7.8) are used to solve for $\{h_j^{t_1-1}, d_i^{t_1-1} j = 1, 2, 3;\ i = 1, \ldots, 4\}$, A_{t_1-1} is formed, and P_{t_1-1} is computed with (7.13). The recursive process is repeated until the h_j^t and d_j^t's have been computed for all $t = t_0, \ldots, t_1$.

Like Prescott and Kydland, we are actually interested in using this algorithm to compute a rational expectations equilibrium for the case of an infinite horizon for the firm, in which case there obtain time invariant laws of motion both for the firm's capital stock and the industry's aggregate capital stock. The firm's problem is to maximize

$$\lim_{t_1 \to \infty} E_{t_0} \sum_{t=t_0}^{t_1} \beta^{(t-t_0)} \left\{ p_t f_0 k_t - w_t k_t - \frac{d}{2} (k_{t+1} - k_t)^2 \right\} \tag{7.15}$$

subject to the laws of motion

$$\begin{aligned} K_{t+1} &= h_0 + h_1 K_t + h_2 w_t + h_3 u_t \\ w_{t+1} &= \lambda_0 + \lambda_1 w_t + \lambda_2 K_t + a_{w_{t+1}} \\ u_{t+1} &= \alpha u_t + a_{u_{t+1}} \end{aligned} \tag{7.16}$$

The solution of this problem for the firm is a linear contingency plan

$$k_{t+1} = d_0 + d_1 k_t + d_2 w_t + d_3 u_t + d_4 K_t, \tag{7.17}$$

which is now time invariant. A *rational expectations equilibrium* is a pair of $\{h_j\}$, $\{d_i\}$, $j = 1,\ 2,\ 3;\ \ i = 1, \ldots, 4$ such that (a) given the h_j's, the d_i's in (7.17) lead to maximization of

(7.15), and (b) the d_i's imply the law of motion for K assumed by firms in their maximization problem, which means that

$$
\begin{aligned}
h_0 &= nd_0 \\
h_1 &= d_1 + nd_4 \\
h_2 &= nd_2 \\
h_3 &= nd_3
\end{aligned}
\tag{7.18}
$$

If the Kydland-Prescott algorithm converges, then it converges to the infinite-horizon, time-invariant equilibrium. That is, set

$$
\begin{aligned}
d_j &= \lim_{t_0 \to -\infty} d_j^{t_0} \qquad & j = 1, \ldots, 4 \\
h_i &= \lim_{t_0 \to -\infty} h_i^{t_0} \qquad & i = 1, \ldots, 3
\end{aligned}
\tag{7.19}
$$

If the limits on the right sides of (7.19) exist, then it can be proved directly that the h_i and d_j defined by (7.19) constitute a rational expectations equilibrium for the infinite horizon setup.

In general, in the presence of feedback from K to w, that is, with λ_2 not zero, there is no guarantee that an infinite horizon equilibrium can be calculated using (7.19). The limits in (7.19) may or may not exist in the presence of feedback from K to w. At present, it is an open question whether, when the limits in (7.19) fail to exist, there still exists an equilibrium for the infinite horizon setup, even though it cannot be calculated by the Kydland-Prescott algorithm.

With $\lambda_2 = 0$ and with $\mid \alpha \mid < 1/\sqrt{\beta}$ and $\mid \lambda_1 \mid < 1/\sqrt{\beta}$, it can be proved that a time invariant equilibrium exists for the infinite horizon setup.

Under conditions delineated by Lucas and Prescott, the rational expectations equilibrium for a model like ours can be computed by solving a particular social planning problem, namely, by maximizing the expected discounted consumer surplus minus the total costs of production. The Kydland-Prescott algorithm is designed for computing rational expectations equilibria in circumstances in which the equivalence between the social planning problem and the competitive equilibrium does not obtain. In our model, the feedback from market-wide capital K to the rental w manifested in equation 7.2 produces an externality that renders the Lucas-Prescott method inapplicable.

3. The Lucas-Prescott Method

An alternative to Kydland-Prescott's algorithm is to represent the feedback from current market wide phenomenona to future w's in a way that preserves the Lucas-Prescott equivalence between the competitive equilibrium and the social planning problem. In the context of the present example, this can be done by introducing feedback directly from u to w, and suppressing the explicited dependence of w on lagged K exhibited by (7.2). This would be accomplished, for example, by replacing (7.2) with

$$w_t = \lambda_0 + \lambda_1 w_{t-1} + \lambda_3 u_{t-1} + a_{wt}. \tag{7.2'}$$

This change restores the equivalence between competitive equilibrium and the social planning problem.

The advantage of the alternative formulation is that by solving the social planning problem, the competitive equilibrium can be calculated more quickly and much more nearly in closed form than by the Kydland-Prescott method. Thus consider the infinite horizon problem, to maximize

$$E_{t_0} \sum_{t=t_0}^{\infty} \beta^{(t-t_0)} \left\{ p_t f_0 k_t - w_t k_t - \frac{d}{2} \left(k_{t+1} - k_t\right)^2 \right\} \tag{7.20}$$

with k_{t_0} given, and subject to

$$p_t = A_0 - A_1 Y_t + u_t \tag{7.21}$$

$$Y_t = n f_0 k_t \tag{7.22}$$

$$\begin{bmatrix} w_t \\ u_t \end{bmatrix} = \begin{bmatrix} c_{11}(L) & c_{12}(L) \\ 0 & c_{22}(L) \end{bmatrix} \begin{bmatrix} v_{1t} \\ v_{2t} \end{bmatrix} = c(L) v_t \tag{7.23}$$

where $c_{ij}(L) = \sum_{k=0}^{\infty} c_{ij,k} L^k$, and all zeroes of *det* $c(z)$ lie outside the unit circle. Here $[v_{1t}, v_{2t},]$ are jointly fundamental for $[w_t, u_t,]$, i.e. they are serially uncorrelated, have means of zero, obey $Ev_{1t}v_{2s} = 0$ for all t and s, and one step ahead linear least squares errors in forecasting (w_t, u_t) by linear functions of lagged w's and u's are linear combinations of v_{1t} and v_{2t}.

Following the procedures in Hansen and Sargent [] and Sargent [], we first solve the certainty version of the above problem, then use the Wiener-Kolmogorov prediction formulas to find the solution under uncertainty. For the certainty version of (7.20), differentiating (7.20) with respect to k_t yields the Euler equation

$$d\beta k_{t+1} - d(1+\beta)k_t + dk_{t-1} = \beta w_t - \beta f_0 p_t$$

At this point, *but not before*, we substitute for p_t from $p_t = A_0 - A_1 n f_0 k_t + u_t$. Substituting at this point and not before is what guarantees that the firm is behaving as a competitor with respect to the output price p_t. Upon substitution we get, after some rearrangement

$$k_{t+1} - \left(1 + \frac{1}{\beta} + \frac{A_1 f_0^2 n}{d}\right) k_t + \frac{1}{\beta} k_{t-1} = \frac{1}{d} w_t - \frac{f_0 A_0}{d} - \frac{f_0}{d} u_t. \tag{7.24}$$

As in Sargent [], it can be shown that this equation itself is the Euler equation associated with the social planning problem, to maximize

$$W_{t_0} = \sum_{t=t_0}^{\infty} \beta^{(t-t_0)} \left\{ [A_0 f_0 n k_t - \frac{1}{2} A_1 (f_0^2 n^2 k_t)^2) + f_0 n u_t k_t] \right.$$
$$\left. - n w_t k_t - \frac{1}{2} n d (k_{t+1} - k_t)^2 \right\}$$

The term in brackets is the area under the demand curve since

$$\int_0^{Y_t} (A_0 - A_1 x + u_t) dx = A_0 Y_t - \frac{1}{2} A_1 Y_t^2 + Y_t u_t$$

The Euler equation (7.24) can be written as

$$(1 - \lambda_1 L)(1 - \lambda_2 L) k_{t+1} = \frac{1}{d} w_t - \frac{f_0 A_0}{d} - \frac{f_0}{d} u_t$$

where $(1 - \lambda_1 L)(1 - \lambda_2 L) = (1 - (1 + \frac{1}{b} + \frac{A_1 f_0^2 n}{d}) L + \frac{1}{b} L^2)$, and where $\beta \lambda_1 = \lambda_2^{-1}$, and where $\lambda_1 < 1 < \frac{1}{\beta} < \lambda_2$. The solution of the Euler equation that satisfies the transversality condition for the firm's and for the social planning problem is (see Sargent [])

$$(1 - \lambda_1 L) k_{t+1} = \frac{-(\lambda_1 \beta) L^{-1}}{1 - (\lambda_1 \beta) L^{-1}} \left\{ \frac{1}{d} w_t - \frac{f_0 A_0}{d} - \frac{f_0}{d} u_t \right\} \tag{7.25}$$

For the problem under uncertainty, the solution is

$$(1 - \lambda_1 L) k_{t+1} = \frac{\lambda_1 \beta}{1 - \lambda_1 \beta} \left(\frac{f_0 A_0}{d} \right)$$
$$+ [-\frac{\lambda_1 \beta}{d} \ \frac{f_0 \beta \lambda_1}{d}] E_t \begin{bmatrix} \frac{L^{-1}}{1 - \lambda_1 \beta L^{-1}} & w_t \\ \frac{L^{-1}}{1 - \lambda_1 \beta L^{-1}} & u_t \end{bmatrix}$$

Given (7.23), we have that

$$(7.27) \qquad E_t \begin{bmatrix} \frac{L^{-1}}{1-\lambda_1\beta L^{-1}} & w_t \\ \frac{L^{-1}}{1-\lambda_1\beta L^{-1}} & u_t \end{bmatrix} = [\frac{L^{-1}}{1-\lambda_1\beta L^{-1}} \; IC(L)] + v_t$$

Using the method of appendix A of Hansen and Sargent [1980], it is readily established that

$$(7.28) \qquad [\frac{L^{-1}}{1-\lambda_1\beta L^{-1}} IC(L)]_+ = \begin{bmatrix} \frac{L^{-1}[C_{11}(L)-C_{11}(\beta\lambda_1)]}{1-\lambda_1\beta L^{-1}} & \frac{L^{-1}[C_{12}(L)-C_{12}(\beta\lambda_1)]}{1-\lambda_1\beta L^{-1}} \\ 0 & \frac{L^{-1}[C_{22}(L)-C_{22}(\beta\lambda_1)]}{1-\lambda_1\beta L^{-1}} \end{bmatrix}$$

Substituting (7.28) and (7.27) into (7.26) gives the equilibrium

$$(7.29) \qquad (1-\lambda_1 L)\, k_{t+1} = \frac{\lambda_1\beta}{1-\lambda_1\beta} \; \frac{f_0 A_0}{d} + \theta_1(L)v_{1t} + \theta_2(L)v_{2t}$$

where

$$\theta_1(L) = \frac{-\lambda_1\beta}{d} \cdot \frac{L^{-1}[C_{11}(L) - C_{11}(\beta\lambda_1)]}{1-\lambda_1\beta L^{-1}}$$

$$(7.30) \qquad \theta_2(L) = \frac{\lambda_1\beta}{d} \; \frac{L^{-1}[C_{12}(L) - C_{12}(\beta\lambda_1)]}{1-\lambda_1\beta L^{-1}}$$

$$+ \frac{f_0\beta\lambda_1}{d} \; \frac{L^{-1}[C_{22}(L) - C_{22}(\beta\lambda_1)]}{1-\lambda_1\beta L^{-1}}$$

Expressions (7.30) can be made to yield explicit formulas for the distributed lag coefficients. Let $\gamma(L) = \sum_{j=0}^{q} \gamma_j L^j$ and $|\,\delta\,| < 1$. Then following the same procedure as in Hansen and Sargent [], it can be proved that

$$(7.31) \qquad L^{-1}[\frac{\gamma(L)-\gamma(\delta)}{1-\delta L^{-1}}] = \gamma_q L^{q-1} + (\gamma_{q-1} + \delta\gamma_q)L^{q-2} + \dots$$

$$+ (\gamma_1 + \delta\gamma_2 + \dots + \delta^{q-1}\gamma_q)L^0).$$

Repeated use of (7.31) in (7.30) converts (7.30) into explicit formulas in terms of lag distributions.

Substituting (7.27) into (7.26) gives

$$(7.32) \qquad (1-\gamma_1 L)k_{t+1} = \frac{\lambda_1\beta}{1-\lambda_1\beta}\frac{f_0 A_0}{d}$$

$$+ [\frac{-\lambda_1\beta}{d}\frac{f_0\beta\lambda_1}{d}]\,[\frac{L^{-1}}{1-\lambda_1\beta L^{-1}} IC(L)]_+ v_t$$

It is useful to consider the case in which $[w_t,\ u_t]$ has an autoregressive representation

$$\begin{bmatrix} A_{11}(L) & A_{12}(L) \\ 0 & A_{22}(L) \end{bmatrix} \begin{bmatrix} w_t \\ u_t \end{bmatrix} = \begin{bmatrix} v_{1t} \\ v_{2t} \end{bmatrix} \tag{7.33}$$

or

$$A(L) \begin{bmatrix} w_t \\ u_t \end{bmatrix} = \begin{bmatrix} v_{1t} \\ v_{2t} \end{bmatrix}$$

where $A(L) = C(L)^{-1}$. Sometimes it will be convenient to parameterize the model in terms of the autoregressive parameters of (7.33). In this case, (7.32) can be written

$$\begin{aligned}(1 - \lambda_1 L)k_{t+1} &= \frac{\lambda_1\beta}{1 - \lambda_1\beta} \frac{f_0 A_0}{d} \\ &+ [-\frac{\lambda_1\beta}{d} \ \frac{f_0\beta\lambda_1}{d}] \ [\frac{L^{-1}}{1 - \lambda_1\beta L^{-1}} I A(L)^{-1}]_+ A(L) \begin{bmatrix} w_t \\ v_t \end{bmatrix}\end{aligned} \tag{7.34}$$

Using

$$C(L) = A(L)^{-1} = \begin{bmatrix} \frac{1}{A_{11}(L)} & \frac{-A_{12}(L)}{A_{11}(L)A_{22}(L)} \\ 0 & \frac{1}{A_{22}(L)} \end{bmatrix},$$

together with (7.34) and (7.28), we obtain the equilibrium in the form

$$(1 - \lambda_1 L)k_{t+1} = \frac{\lambda_1\beta}{1 - \lambda_1\beta} \frac{f_0 A_0}{d} + [-\frac{\lambda_1\beta}{d} \frac{f_0\beta\lambda_1}{d}] \tag{7.35}$$

$$\cdot \begin{bmatrix} \left\{\frac{L^{-1}}{1-\beta\lambda_1 L^{-1}} [\frac{1}{A_{11}(L)} - \frac{1}{A_{11}(\lambda_1\beta)}]\right\} & \left\{\frac{-L^{-1}}{1-\beta\lambda_1 L^{-1}} [\frac{A_{12}(L)}{A_{11}(L)A_{22}(L)} - \frac{A_{12}(\lambda_1\beta)}{A_{11}(\lambda_1\beta)A_{22}(\lambda_1\beta)}]\right\} \\ 0 & \frac{L^{-1}}{1-\beta\lambda_1 L^{-1}} [\frac{1}{A_{22}(L)} - \frac{1}{A^{22}(\lambda_1\beta)}] \end{bmatrix}$$

$$\cdot \begin{bmatrix} A_{11}(L) & A_{12}(L) \\ 0 & A_{22}(L) \end{bmatrix} \begin{bmatrix} w_t \\ u_t \end{bmatrix}$$

Performing the indicated matrix multiplications gives

$$k_{t+1} = \lambda_1 k_t + \frac{\lambda_1\beta}{1 - \lambda_1\beta} \frac{f_0 A_0}{d} + b_1(L)w_t + b_2(L)u_t$$

where

$$\begin{aligned} b_1(L) &= -\frac{\lambda_1\beta}{d}\left\{L^{-1}[I - \frac{A_{11}(L)}{A_{11}(L)(\lambda_1\beta)} \frac{1}{1 - \lambda_1\beta L^{-1}}\right\} \\ b_2(L) &= \frac{\lambda_1\beta}{d}\left\{L^{-1}[\frac{A_{12}(L)A_{22}(\lambda_1\beta) - A_{12}(\lambda_1\beta)A_{22}(L)}{A_{11}(\lambda_1\beta)A_{22}(\lambda_1\beta)}] \cdot \frac{1}{1 - \lambda_1\beta L^{-1}}\right\} \\ &+ \frac{\lambda_1\beta f_0}{d}\left\{L^{-1}[I - \frac{A_{22}(L)}{A_{22}(\lambda_1\beta)}] \frac{1}{1 - \lambda_1 B L^{-1}}\right\}. \end{aligned}$$

4. A Corn-Hog Model

This section describes a simplified model of the "corn-hog cycle". We have adopted the most rudimentary specification of technologies, retaining only those elements that are essential to exhibit what we believe are the key features such a model must have. A more realistic specification of the technologies would involve more state variables and more complicated dynamics, but would not involve any essential analytical complications.

There are m identical corn farmers, each of whom maximizes

$$E_{t_0} \sum_{t=t_0}^{\infty} \beta^{(t-t_0)} \left\{ p_{c_t} c_t - w_t k_{c_t} - \frac{d}{2}(k_{ct+1} - k_{ct})^2 \right\} \tag{7.36}$$

where p_{ct} is the price of corn, c_t is output of corn, k_{c_t} is the capital stock of the corn producers, and w_t is the rental rate on capital. Output of corn obeys

$$c_t = f k_{c_t}, \quad f > 0 \tag{7.37}$$

The corn producer faces the stochastic processes for p_{ct} and w_t as a price taker. The rental is assumed to be the first element of a $(p_w \times 1)$ stochastic process z_t which obeys the q^{th} order autoregressive law

$$\{I - \rho_1 L - \ldots - \rho_q L^q\} z_t = v_t^z \quad \text{or} \quad \rho(L) z_t = v_t^z \tag{7.38}$$

where the zeroes of $det\,\{I - \rho_1 z - \ldots - \rho_q z^q\}$ lie outside the unit circle. Here v_t^z is a serially uncorrelated vector process with mean zero. We assume that v_t^z is fundamental for z_t. The assumptions about the stochastic process for p_{ct} will be filled in later. .

The hog industry consists of n identical producers each of whom maximizes

$$E_{t_0} \sum_{t=t_0}^{\infty} \beta^{(t-t_0)} \left\{ p_{ht} h_t - r_t k_{ht} - \frac{e}{2}(k_{ht+1} - k_{ht})^2 - p_{ct} c_{ht} \right\} \tag{3.4}$$

where p_{ht} is the price of hogs, c_{ht} the consumption of corn by hogs, k_{ht} the number of hogs, h_t sales of hogs, $r_r k_{ht}$ is miscellaneous expenses to maintain k_{ht} hogs. The technology is assumed to be

$$\begin{aligned} c_{ht} &= \gamma k_{ht} & \gamma > 0 \\ h_t &= (1 + \phi) k_{ht} - k_{ht+1}, & \phi > 0 \end{aligned} \tag{7.40}$$

where ϕ is governed by the reproduction rate of pigs, which is assumed exogenous here but would be a decision variable in a more realistic analysis. In (7.39), the term $(e/2)(k_{ht+1} - k_{ht})^2$ represents costs of adjusting the number of pigs.

The price r_t is the first element of a $(p_r \times 1)$ vector x_t which follows the q^{th} order autoregressive process

$$(I - \delta_1 L - \delta_2 L^2 - \cdots - \delta_q L^q)x_t = v_t^x$$

or

$$\delta(L)x_t = v_t^x \ , \tag{7.41}$$

where the zeroes of $det\zeta(z)$ lie outside the unit circle and v_t^x is a fundamental white noise vector for x_t. The demand for hogs is given by

$$p_{ht} = A_0 - A_1 H_t + u_{ht} \ , \quad A_0, \ A_1 > 0 \tag{7.42}$$

where $H_t = nh_t$, and where u_{ht} is a stochastic shock to demand that obeys the autoregressive law

$$\{1 - \alpha_1 L - \ldots - \alpha_s L^2\}u_{ht} = v_t^h$$

or

$$\alpha(L)u_{ht} = v_t^h \tag{7.43}$$

where the zeroes of $\alpha(z)$ lie outside the unit circle and v_t^h is the fundamental white noise for u_{ht}.

The demand for corn is the sum of the demand derived from hog production, $C_{ht} = nc_{ht}$ and the demand for final consumption, C_{ct}. The demand for final consumption obeys

$$C_{ct} = \beta_0 - \beta_1 p_{ct} + u_{ct} \quad \beta_0 > 0, \beta_1 > 0, \tag{7.44}$$

where u_{ct} is a demand shock that obeys the autoregressive law

$$\{1 - \gamma_1 L - \ldots - \gamma_s L^s\}u_{ct} = v_t^c$$

or

$$\gamma(L)u_{ct} = v_t^c \tag{7.45}$$

where the zeroes of $\gamma(z)$ lie outside the unit circle, and v_t^c is the fundamental white noise for u_{ct}. The total hog-derived demand for corn from (7.39) is

$$C_{ht} = \gamma K_{ht}$$

where $K_{ht} = nk_{nt}$ and $K_{ct} = mk_{ct}$. The equilibrium condition in the market for corn is

$$C_{ct} + C_{ht} = fK_{ct}$$

or

$$\beta_0 - \beta_1 p_{ct} + u_{ct} + \gamma K_{ht} = fK_{ct}$$

which implies

$$p_{ct} = \frac{1}{\beta_1}[\gamma K_{ht} + \beta_0 + u_{ct} - fK_{ct}]. \tag{7.46}$$

At this point it is convenient to define the vectors

$$\begin{aligned} \bar{z}_t &= [z_t', z_{t-1}', \ldots, z_{t-q+1}']' \\ \bar{x}_t &= [x_t', x_{t-1}', \ldots, x_{t-q+1}']' \\ \bar{u}_{ht} &= [u_{ht-1}, \ldots, u_{ht-s+1}]' \\ \bar{u}_{ct} &= [u_{ct}, u_{ct-1} \ldots, u_{ct-s+1}]'. \end{aligned}$$

The farmers in each market need to form expectations about future prices of corn and hogs in order to solve the maximum problems (7.36) and (7.39). Since future corn and hog prices will depend on future state variables in each market, including the capital stocks in each market, farmers in each market need to form a view about the laws of motion of the market wide stocks of capital in both markets. We assume that farmers views about these laws of motion are correct. It will turn out that the laws of motion for the market wide capital stocks in the two industries will have the forms

$$K_{ct+1} = G_c(\bar{z}_t, \bar{x}_t, \bar{u}_{ht}, K_{ct}, K_{ht}, 1) \tag{7.47}$$

$$K_{ht+1} = G_h(\bar{z}_t, \bar{x}_t, u_{ht}u_{ct}, K_{ct}, K_{ht}, 1) \tag{7.48}$$

where both G_c and G_n are linear functions. Notice that (7.47) and (7.48) share a common set of arguments, so that any state variable for one market is also a state variable for the other.

We are now in a position to state well posed optimum problems for firms in each industry. Firms in the hog industry maximize

$$\begin{aligned} E_{t_0} \sum_{t=t_0}^{\infty} \beta^{(t-t_0)} \Big\{ & [A_0 - A_1((1+\phi)K_{ht} - K_{ht+1}) + u_{ht}] \cdot [(1+\phi)k_{ht} - k_{ht+1}] \\ & - r_t k_{ht} - \frac{e}{2}(k_{ht+1} - k_{ht})^2 \\ & - \gamma k_{ht} \cdot [\frac{1}{\beta_1}(\gamma K_{ht} + \beta_0 + u_{ct} - f K_{ct})] \Big\} \end{aligned} \tag{7.49}$$

Firms in the corn industry maximize

$$\begin{aligned} E_{t_0} \sum_{t=t_0}^{\infty} \beta^{(t-t_0)} \{ & [\frac{1}{\beta_1}(\gamma K_{ht} + \beta_0 + u_{ct} - f K_{ct})] f k_{ct} \\ & - w_t k_{ct} - \frac{d}{2}(k_{ct+1} - k_{ct})^2 \} \end{aligned} \tag{7.50}$$

Here it should be noted that $p_{ct} = \frac{1}{\beta_1}(\gamma K_{ht} + \beta_0 + u_{ct} - f K_{ct})$ and $p_{ht} = A_0 - A_1[(1+\phi)K_{nt} - K_{nt+1}] + u_{ht}$, and that these expressions for p_{ct} and p_{ht} have been substituted into (7.36) and (7.39) to obtain (7.50) and (7.49), respectively. The maximization in (7.49) and (7.50) is subject to the following laws of motion, which the firms in each industry take as given:

$$\zeta(L) z_t = v_t^z \tag{7.38}$$

$$\delta(L) x_t = v_t^x \tag{7.41}$$

$$\alpha(L) u_{nt} = v_t^h \tag{7.43}$$

$$\gamma(L) u_{ct} = v_t^c \tag{7.45}$$

$$K_{ct+1} = G_c(\bar{z}_t, \bar{x}_t, \bar{u}_{ht}, \bar{u}_{ct}, K_{ct}, K_{ht}, 1) \tag{7.47}$$

$$K_{nt+1} = G_n(\bar{z}_t, \bar{x}_t, \bar{u}_{nt}, K_{ct}, K_{nt}, 1) \tag{7.48}$$

The optimizations in (7.49) and (7.50), respectively, are over linear contingency plans of the forms

$$k_{ht+1} = g_h(\bar{z}_t, \bar{x}_t, \bar{u}_{ht}, \bar{u}_{ct}, K_{ct}, K_{ht}, 1, k_{ht} \tag{7.51}$$

$$k_{ct+1} = g_c(\bar{z}_t, \bar{x}_t, \bar{u}_{ht}, \bar{u}_{ct}, K_{ct}, K_{ht}, 1, k_{ct}) \tag{7.52}$$

where both g_h and g_c are linear functions.

We are now in a position to define a rational expectations equilibrium for this pair of industries.

Definition 7.2: A *rational expectations equilibrium* is four linear functions (7.47), (7.48), (7.51), and (7.52) such that

(*a*) Given the aggregate laws of motion (7.47) and (7.48), the contingency plans (7.51) and (7.52) maximize the expected present values, (7.49) and (7.50), respectively.

(*b*) The contingency plans of the representative firms in each industry (7.51) and (7.52) imply the aggregate laws of motion (7.47) and (7.48), so that

$$\begin{aligned} &G_c(\bar{z}_t, \bar{x}_t, \bar{U}_{ht}, \bar{U}_{ct}, K_{ct}, K_{nt}, 1) = \\ &\qquad mg_c(\bar{z}_t, \bar{x}_t, \bar{U}_{nt}, \bar{U}_{ct}, K_{ct}, K_{nt}, 1, k_{ct}) \end{aligned} \tag{7.53}$$

$$\begin{aligned} &G_h(\bar{z}_t, \bar{x}_t, \bar{u}_{ht}, \bar{u}_{ct}, K_{ct}, K_{nt}, 1) \\ &\qquad ng_h(\bar{z}_t, \bar{x}_t, \bar{u}_{nt}, \bar{u}_{ct}, K_{ct}, K_{nt}, 1, k_{nt}). \end{aligned} \tag{7.54}$$

Let us indicate how the Kydland-Prescott algorithm can be used to compute the equilibrium. Write (7.38), (7.41), (7.43) and (7.44) as

$$\begin{aligned} \bar{z}_{t+1} &= \bar{\rho}\bar{z}_t + v^z_{t+1} \\ \bar{x}_{t+1} &= \bar{\sigma}\bar{x}_t + v^x_{t+1} \\ \bar{u}_{ht+1} &= \bar{\alpha}\bar{u}_{ht} + v^h_{t+1} \\ \bar{u}_{ct+1} &= \bar{\gamma}\bar{u}_{ct} + v^c_{t+1} \end{aligned}$$

where $\bar{\rho} = [\rho_1\rho_2 \ldots \rho_q], \bar{\sigma} = [\sigma_1\sigma_2 \ldots \sigma_q], \bar{\alpha} = [\alpha_1\alpha_2 \ldots \alpha_s], \bar{\gamma} = [\gamma_1\gamma_2 \cdots \gamma_s]$. Further, let G_{hi} and G_{cj} be the partial derivatives of (7.48) and (7.47), respectively, with respect to their i^{th}

and j^{th} arguments. Then from the point of view of the corn industry, the state transition equation is

$$
\begin{bmatrix} \bar{z}_{t+1} \\ \bar{x}_{t+1} \\ \bar{u}_{ht+1} \\ \bar{u}_{ct+1} \\ K_{ct+1} \\ K_{ht+1} \\ 1 \\ k_{ct+1} \end{bmatrix}
=
\begin{bmatrix}
\bar{\rho} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \bar{\sigma} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \bar{\alpha} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \bar{\gamma} & 0 & 0 & 0 & 0 \\
G_{c1}^t & G_{c2}^t & G_{c3}^t & G_{c4}^t & G_{c5}^t & G_{c6}^t & G_{c7}^t & G_{c8}^t \\
G_{h1}^t & G_{h2}^t & G_{h3}^t & G_{h4}^t & G_{h5}^t & G_{h6}^t & G_{h7}^t & G_{h8}^t \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix} \bar{z}_t \\ \bar{x}_t \\ \bar{u}_{ht} \\ \bar{u}_{ct} \\ K_{ct} \\ K_{ht} \\ 1 \\ k_{ct} \end{bmatrix}
+
\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} v_{ct}
+
\begin{bmatrix} v_{t+1}^z \\ v_{t+1}^x \\ v_{t+1}^n \\ v_{t+1}^c \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}
$$

or

$$X_{ct+1} = A_{ct} X_{ct} + B_c v_{ct} + \varepsilon_{ct+1} \tag{7.55}$$

For the representative hog producer the state transition equation is

$$
\begin{bmatrix} \bar{z}_{t+1} \\ \bar{x}_{t+1} \\ u_{ht+1} \\ \bar{u}_{ct+1} \\ K_{ct+1} \\ K_{nt+1} \\ 1 \end{bmatrix}
=
\begin{bmatrix}
\bar{\rho} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \bar{\sigma} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \bar{\alpha} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \bar{\gamma} & 0 & 0 & 0 & 0 \\
G_{c1}^t & G_{c2}^t & G_{c3}^t & G_{c4}^t & G_{c5}^t & G_{c6}^t & G_{c7}^t & G_{c8}^t \\
G_{h1}^t & G_{h2}^t & G_{h3}^t & G_{h4}^t & G_{h5}^t & G_{h6}^t & G_{h7}^t & G_{h8}^t \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix} \bar{z}_t \\ \bar{x}_t \\ \bar{u}_{ht} \\ \bar{u}_{ct} \\ K_{ct} \\ K_{nt} \\ 1 \\ k_{ht} \end{bmatrix}
+
\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} v_{ht}
+
\begin{bmatrix} v_{t+1}^z \\ v_{t+1}^x \\ v_{t+1}^h \\ v_{t+1}^c \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}
$$

or

$$X_{ht+1} = A_{ht} X_{ht} + B_h v_{ht} + \varepsilon_{ht+1} \tag{7.56}$$

Now the corn producer's problem (7.50) can be expressed as the maximization of

$$(7.50)' \qquad E_{t_0} \sum_{t=t_0}^{\infty} \beta^{(t-t_0)} \{X_{ct}^T R_c X_{ct} + v_{ct}^T Q_c v_{ct}\}$$

subject to

$$X_{ct+1} = A_{ct} X_{ct} + B_c v_{ct} + \varepsilon_{ct+1}$$

where R_c and Q_c are matrices conformable with X_{ct} which make (7.50)′ equivalent to (7.50)′. Similarily the hog producers' problem (7.49) can be expressed as the maximization of

$$(7.49)' \qquad E_{t_0} \sum_{t=t_0}^{\infty} \beta^{(t-t_0)} \{X_{ht}^T R_h X_{ht} + v_{ht}^T Q_h v_{ht}\}$$

subject to

$$X_{ht+1} = A_{ht} X_{ht} + B_h v_{ht} + \varepsilon_{ht+1}$$

The Kydland-Prescott algorithm can be used simultaneously to compute the parameters of G_c^t and G_h^t that appear in A_{ct} and A_{ht}, as well as the optimum decision rules g_c^t and g_h^t. The equations for the optimum decision rules can be written, as in section 1:

$$(7.57) \qquad \begin{aligned} -[-g_{h1}^t, -g_{h2}^t, -g_{h3}^t, -g_{h4}^t, -g_{h5}^t, -g_{h6}^t, -g_{h7}^t, (1-g_{h8}^t)] = \\ \beta(\beta B_h' P_{ht+1} B_h + Q_h)^{-1} A_{ht} \end{aligned}$$

$$(7.58) \qquad \begin{aligned} -[-g_{c1}^t, -g_{c2}^t, -g_{c3}^t, -g_{c4}^t, -g_{c5}^t, -g_{c6}^t, -g_{c7}^t, (1-g_{c8}^t)] = \\ \beta(\beta B_c' P_{ct+1} B_c + Q_c)^{-1} A_{ct} \end{aligned}$$

where P_{ct} and P_{ht} are obtained from the matrix Riccati difference equations

$$(7.59) \qquad \begin{aligned} P_{ct} &= \beta A_{ct}' P_{ct+1} A_{ct} + R_{ct} \\ &- \beta^2 A_{ct}' P_{ct} B_c [\beta B_c' P_{ct+1} B_c + Q_c]^{-1} B_c' P_{ct+1} A_{ct} \end{aligned}$$

$$(7.60) \qquad \begin{aligned} P_{ht} &= \beta A_{ht}' P_{ht+1} A_{ht} + R_{ht} \\ &- \beta^2 A_{ht}' P_{ht} B_h [\beta B_h' P_{ht+1} B_n + Q_h]^{-1} B\prime_h P_{ht+1} A_{ht} \end{aligned}$$

starting from $P_{ct_1+1} = 0, \quad P_{ht_1+1} = 0$.

The equilibrium conditions (7.53) and (7.54) supply us with the linear equations

$$(7.61) \qquad \begin{aligned} &G_{c_1}^t = mg_{c_1}^t, \quad G_{c_2}^t = mg_{c_2}^t, \quad G_{c_3}^t = mg_{c_3}^t, \quad G_{c_4}^t = mg_{c_4}^t \\ &G_{c_5}^t = mg_{c_5}^t + g_{c_8}^t, \quad G_{c_6}^t = mg_{c_6}^t, \quad G_{c_7}^t = mg_{c_7}^t \end{aligned}$$

$$G^t_{h_1} = ng^t_{h_1},\quad G^t_{h_2} = ng^t_{h_2},\quad G^t_{h_3} = ng^t_{h_3},\quad G^t_{h_4} = ng^t_{h_4}$$
$$(7.62)\qquad G^t_{h_5} = ng^t_{h_5},\quad G^t_{h_6} = ng^t_{h_6} + g^t_{h_8},\quad G^t_{h_6} = ng^t_{h_6},\quad G^t_{h_7} = ng^t_{h_7}$$

The reader can verify that equations (7.57), (7.58), (7.61) and (7.62) are $2(2s+2q+3)+2(2s+2q+4)$ linear equations in the same number of unknowns, where the unknowns are the $g^t_{h_j}, g^t_{c_j}, G^t_{h_i} G^t_{c_i}$, for $j=1,\dots,8, i=1,\dots,7$.

As in section 1, the computation strategy is to solve (7.57), (7.58), (7.61), and (7.62) jointly, starting from $P_{ct_1+1}=0, P_{ht+1}=0$. Equations (7.59) and (7.60) are used to "back-date" P_{ct} and P_{ht}. The idea is to iterate on these equations and take the limits of the $G^{t_0}_{h_j}, G^{t_0}_{c_j}, g^{t_0}_{h_j}, g^{t_0}_{c_j}$'s as $t_0 \to -\infty$.

5. Solving the Corn Hog Model a la Lucas-Prescott

This section shows how to calculate the equilibrium of the corn-hog model by using the methods of section 2. As in section 2, the idea is it obtain the Euler equations for the representative firms' problems, then to substitute into those Euler equations expressions for equilibrium prices in terms of market wide stocks of factors, and then finally to solve the resulting system of difference equations subject to the transversality conditions for firms' problems.

Taking the hog producer's problem first, recall that the hog producers problem is to maximize

$$(7.63)\qquad E_{t_0}\sum_{t=t_0}^{\infty}\beta^{(t-t_0)}\Big\{P_{ht}[(1+\phi)k_{ht}-k_{ht+1}]-r_tk_{ht}-\frac{e}{2}\,(k_{ht+1}-k_{ht})^2-\rho_{ct}\gamma k_{ht}\Big\}.$$

As in Sargent [], we solve this problem by first solving the problem assuming there is not uncertainty. From the certain version of the problem the Euler equation is

$$(7.64)\qquad \begin{aligned}&\beta ek_{ht+1}-(1+\beta)\,ek_{ht}+ek_{ht-1}\\ &\qquad = -\beta(1+\phi)p_{ht}+p_{ht-1}+\beta\gamma p_{ct}+\beta\gamma_t\end{aligned}$$

From the demand curve for hogs we have

$$p_{ht} = A_0 - A_1[(1+\phi)K_{ht}-K_{ht+1}]+u_{ht}$$

so that

$$
\begin{aligned}
p_{ht-1} - \beta(1+\phi)p_{ht} &= \{A_0 - \beta(1+\phi)A_0\} + u_{ht-1} - \beta(1+\phi)u_{ht} \\
&+ \{\beta(1+\phi)^2 A_1 + A_1\}K_{ht} - \beta(1+\phi)A_1 + K_{ht+1} \\
&- A_1(1+\phi)K_{ht-1}
\end{aligned} \tag{7.65}
$$

We also have from the demand curve for corn

$$
p_{ct} = \frac{1}{\beta_1}[\gamma K_{ht} + B_0 + u_{ct} - fK_{ct}]. \tag{7.66}
$$

Substituting (7.65) and (7.66) into (7.64) and using $K_{ht} = nk_{ht}$ and $K_{ct} = mk_{ct}$ gives

$$
\begin{aligned}
&\{\beta e + \beta n(1+\phi)A_1\}nk_{ht+1} - \left[(1+\beta)e + n\{\beta(1+\phi)^2 A_1 + A_1\} + \frac{n\beta\gamma^2}{\beta_1}\right]nk_{ht} \\
&+ \{e + n(1+\phi)A_1\}nk_{ht-1} + \frac{\beta\gamma f}{\beta_1}nmk_{c_t} \\
&= nu_{ht-1} - n\beta(1+\phi)u_{ht} + \beta nr_t + \frac{n\beta\gamma}{\beta_1}u_{c_t} \\
&+ \frac{n\beta\gamma\beta_0}{\beta_1} + n[A_0 - \beta(1+\phi)A_0]
\end{aligned} \tag{7.67}
$$

Equation (7.67) is one of a pair of Euler-like equations whose solution will determine a competitive equilibrium for the two industries.

Corn farmers maximize their expected present value

$$
E_{t_0}\sum_{t=t_0}^{\infty} \beta^{(t-t_0)}\ \{p_{ct}fk_{ct} - w_t k_{ct} - \frac{d}{2}(k_{ct+1} - k_{ct})^2\}.
$$

Again following the procedure in Sargent, we first consider the version of this problem under certainty. Under certainty the Euler equation is

$$
\beta dk_{ct+1} - d(1+\beta)k_{ct} + dk_{ct-1} = \beta w_t - \beta f p_{ct}
$$

Substituting (7.66) into the above equation gives

$$
\begin{aligned}
&\frac{\beta f\gamma}{\beta_1}mnk_{ht} + \beta dmk_{ct+1} - [d(1+\beta) + \frac{m\beta f^2}{\beta_1}]mk_{c_t} + mdk_{ct-1} \\
&= \beta mw_t - \frac{m\beta f}{\beta_1}u_{ct} - \frac{m\beta f\beta_0}{\beta_1}
\end{aligned} \tag{7.68}
$$

Equations (7.67) and (7.68) are in the form of a pair of Euler equations for k_{ht} and k_{ct}. It is reasonable to pose the integrability question: for what optimum problem are these Euler equations first order necessary conditions? The answer is not surprising, given the results of Lucas and Prescott []. Notice that since the demand curve for final consumption of corn is $c_{c_t} = \beta_0 - \beta_1 p_{c_t} + u_{c_t}$, the area under the demand curve for final consumption of corn is

$$\int_0^{C_{ct}} \frac{1}{\beta_1}[\beta_0 + u_{ct} - x]dx$$
$$= \frac{\beta_0}{\beta_1}C_{c_t} + \frac{1}{\beta_1}u_{ct}C_{ct} - \frac{1}{2\beta_1}C_{c_t}^2.$$

Substituting for C_{ct} from the equilibrium condition $C_{ct} = fK_{ct} - \gamma K_{ht}$ gives the following formula for the area under the demand curve for corn:

$$\begin{aligned} &\frac{\beta_0}{\beta_1}[fK_{ct} - \gamma K_{ht}] + \frac{1}{\beta_1}u_{ct}[fK_{ct} - \gamma K_{ht}] \\ &\quad - \frac{1}{2\beta_1}fK_{c_t} - \gamma K_{ht}]^2 \end{aligned} \tag{7.69}$$

The area under the demand curve for hogs is:

$$\begin{aligned} &\int_0^{H_t}(A_0 - A_1 x + U_{ht})dx \\ &= A_0 H_t - \frac{1}{2}A_1 H_h^2 + H_t u_t \\ &= A_0[(1+\phi)K_{nt} - K_{nt+1}] \\ &- \frac{1}{2}A_1[(1+\phi)K_{ht} - K_{nt+1}]^2 \\ &+ [(1+\phi)K_{nt} - K_{nt+1}]u_{ht} \end{aligned} \tag{7.70}$$

Using (7.69) and (7.70), consider the following social planning problem: to maximize

$$\begin{aligned} E_{t_0}\sum_{t=t_0}^{\infty}\beta^{t-t_0}\{\{&A_0[(1+\phi)nk_{ht} - nkh_{t+1}] - \frac{1}{2}A_1[(1+\phi)nk_{ht} - nk_{ht+1}]^2 \\ &+ [(1+\phi)nk_{ht} - nk_{ht+1}]u_{ht}\} \\ &+ \{\frac{\beta_0}{\beta_1}[fmk_{c_t} - \gamma nk_{ht}] + \frac{1}{\beta_1}u_{ct}[fmk_{c_t} - \gamma nk_{ht}] \\ &- \frac{1}{2\beta_1}[fmk_{c_t} - \gamma_n k_{ht}]^2 \\ &- w_t mk_{ct} - r_t nk_{ht} - \frac{d}{2}m(k_{ct+1} - k_{ct})^2 \\ &- \frac{e}{2}n(k_{ht+1} - k_{nt})^2\} \end{aligned}$$

where the maximization is subject to the given stochastic processes (7.38), (7.41), (7.43), and (7.45) for w_t, r_t, u_{ct}, and u_{ht}, and the information set $\{k_{ct}, k_{ht}, \bar{z}_t, \bar{x}_t, \bar{u}_{ct}, \bar{u}_{ht},\}$. The maximization is over contingency plans setting $\{k_{ht+1}, k_{ct+1},\}$ as a linear function of this information set.

This social planning problem amounts to maximizing the expected discounted sum of the area under the demand curves for hogs and final consumption of corn *minus* the total social costs of production. It is straightforward to verify that the solution of this social planning problem is equivalent with the competitive equilibrium law of motion for $\{k_{ht+1}, k_{ct+1},\}$. The proof of this claim can be obtained by first obtaining the Euler equations for the social planning problem, and noting that they are exactly the two difference equations (7.67) and (7.68) whose solutions determine the competitive equilibrium. Next, it can be verified that the transversality condition for the social planning problem enforces the same solution of (7.67) and (7.68) as does the transversality conditions for the representative firms' problems.

The Euler equations (7.67) and (7.68) can be written

$$\{\beta L^{-1} G_{-1} + G_0 + G_1 L\} k_t = H_0 + H_1(L) b_t \tag{7.71}$$

where

$$k_t = \begin{bmatrix} k_{ht} \\ k_{ct} \end{bmatrix}, b_t = \begin{bmatrix} r_t \\ w_t \\ u_{ht} \\ u_{ct} \end{bmatrix}$$

$$H_0 = \begin{bmatrix} \frac{n\beta\gamma\beta_0}{\beta_1} + n[A_0 - \beta(1+\phi)A_0] \\ -\frac{m\beta f}{\beta_0} \end{bmatrix}$$

$$H_1(L) = \begin{bmatrix} \beta_n & 0 & nL - \beta_n(1+\phi) & \frac{n\beta\gamma}{\beta_1} \\ 0 & \beta m & 0 & \frac{-m\beta f}{\beta_1} \end{bmatrix}.$$

$$G_1 = G_{-1}^T = \begin{bmatrix} en + n^2(1+\phi)A_1 & 0 \\ 0 & dm \end{bmatrix}$$

$$G_0 = \begin{bmatrix} -[(1+\beta)e + n\{\beta(1+\phi)^2 A_1 + A_1 + \frac{n\beta\gamma^2}{\beta_1}]n & \frac{\beta\gamma f}{\beta_1} nm \\ \frac{\beta\gamma f}{\beta_1} mn & -[d(1+\beta) + \frac{m\beta f^2}{\beta_1}]m \end{bmatrix}$$

Methods for solving matrix Euler equations like (7.71) subject to boundary conditions are described by Hansen and Sargent [1981].

Chapter 8

Combining Recursive Optimization and Classical Filtering to Compute Solutions of Control Problems

1. Introduction

This chapter describes a method for solving linear quadratic optimal control problems of a kind that often arise in linear "rational expectations" models. The method is a variant of one proposed by Hansen and Sargent [1981]. The idea behind the method is to use recursive methods to factor the matrix lag operator polynominal that appears in the Euler equations, and to use classical Wiener-Kolmogorov filtering formulas to compute the "feedforward" part of the optimal control. We have already seen many examples of models that fit into the control problem studied here.

2. The Problem

Consider the problem, maximize

$$E_{t_0} \lim_{t_1 \to \infty} \frac{1}{(t_1 - t_0)} \sum_{t=t_0}^{t_1} (x_t' R x_t + v_t' Q v_t - 2a_t' R_{21}' v_t) \tag{8.1}$$

subject to

$$x_{t+1} = Ax_t + Bv_t \tag{8.2}$$

and x_{t_0} given, where a_t is a set of components of a $(p \times 1)$ vector z_t governed by the r^{th} order autoregressive process

$$z_t = \rho_1 z_{t-1} + \cdots + \rho_r z_{t-r} + v_t^z$$

or

$$\rho(L) z_t = v_t^z \tag{8.3}$$

The ρ_j are matrices conformable to z_t. We assume that the zeroes of $det\, \rho(z)$ are outside the unit circle, and that v_{t+1}^z is orthogonal to $\{z_t, z_{t-1}, \ldots\}$. Here x_t is an $(n \times 1)$ vector of *states*, v_t a $(k \times 1)$ vector of *controls*, A an $(n \times n)$ matrix, and B an $n \times k$ matrix. At time

t, the planner is assumed to know $\{x_t, z_t, z_{t-1}, \ldots\}$ and is supposed to set the control v_t as a linear function of these variables. The matrix Q is negative definite; the matrix R_{21} is not restricted; the matrix R is negative semi-definite.

We are interested in instances in which n is small relative to pr. In such instances, there is an advantage to solving (8.1)–(8.3) using a method that takes into account the features of the problem as a special case of the general linear optimal regulator problem. The method that we shall use is a mixture of dynamic programming and discrete time calculus of variations methods.

To begin, we first consider the related problem of maximizing

$$\lim_{t_1 \to \infty} \frac{1}{(t_1 - t_0)} \sum_{t=t_0}^{t_1} (x_t' R x_t + v_t' Q v_t) \tag{8.4}$$

subject to $x_{t+1} = Ax_t + Bv_t$ with x_{t_0} given. The solution of this problem is a linear feedback rule

$$v_t = -F x_t \tag{8.5}$$

where

$$F' = A'PB[B'PB + Q]^{-1} \tag{8.6}$$

and where P is the negative definite solution of the algebraic Riccati equation

$$P = A'PA + R - A'PB(B'PB + Q)^{-1} B'PA. \tag{8.7}$$

Under the assumption that the pair (A, B) is stabilizable, the unique negative definite solution of (8.7) is the limit point of iterations on the matrix Riccati difference equation,

$$P_{t-1} = A'P_tA + R - A'P_tB(B'P_tB + Q)^{-1}B'P_tA$$

as $t \to -\infty$, starting from $P_{t_1} \equiv 0$. Let $-R = G^T G$, and assume that the pair (A, G) is *detectable*. Then, under the assumption that $[A, B]$ is stabilizable, the closed loop system, derived by substituting $v_t = -Fx_t$ into $x_{t+1} = Ax_t + Bv_t$, namely,

$$x_{t+1} = (A - BF)x_t$$

is asymptotically stable, meaning that the eigenvalues of $(A - BF)$ are less than unity in modulus. We assume that the pair (A, B) is *controllable*, implying that it is stabilizable.

We now return to the problem (8.1)-(8.3). First use (8.3) to write

$$(L^{-1}I - A)x_t = Bv_t$$

or

$$x_t = (L^{-1}I - A)^{-1}Bv_t \tag{8.8}$$

Equation (8.8) needs to be interpreted carefully, since the eigenvalues of A have not been restricted directly, and since the infinite sum $(L^{-1}I - A)^{-1} = L\{I + AL + A^2L^2 + \ldots\}$ is not convergent if an eigenvalue of A exceeds unity in modulus. Nevertheless, if v_t has behaved suitably in the past, it is appropriate to regard (8.8) as giving

$$x_t = \sum_{j=0}^{\infty} A^j Bv_{t-j-1}, \tag{8.8'}$$

where by "suitably" we mean in such a manner as to guarantee convergence of the sum. Since we have assumed that (A, B) is controllable, it is permissable for us to think of the system as having arrived at its arbitrary initial state x_{t_0} via the application of an appropriate sequence of controls in the past. (A version of this argument would also work and we only assumed that (A, B) was stabilizable.) The preceding interpretation is not the only one that would validate our procedures, but it is an acceptable one.

Substituting (8.8) for x_t in (8.1) gives the expression for the objective function

$$E_{t_0} \lim_{t_1 \to \infty} \frac{1}{(t_1 - t_0)} \sum_{t=t_0}^{t_1} \Big\{[(L^{-1}I - A)^{-1}Bv_t]' R[(L^{-1}I - A)^{-1}Bv_t] + v_t'Qv_t - 2a_t'R_{21}'v_t\Big\} \tag{8.9}$$

We shall solve the problem of maximizing (8.9) over rules for v_t by using the certainty equivalence principle. First, we solve the certainty problem to maximize

$$\lim_{t_1 \to \infty} \frac{1}{(t_1 - t_0)} \sum_{t=t_0}^{t_1} \Big\{[(L^{-1}I - A)^{-1}Bv_t]' R[(L^{-1}I - A)^{-1}Bv_t] + v_t'Qv_t - 2a_t'R_{21}'v_t\Big\} \tag{8.10}$$

where $\{a_t, t_0 \leq t \leq t_1\}$ is regarded as a bounded sequence. By differentiating with respect to successive v_t's one obtains the system of Euler equations

$$\{B'(LI - A')^{-1}R(L^{-1}I - A)^{-1}B + Q\}v_t = R_{21}a_t \tag{8.11}$$

In addition to these Euler equations, there is a set of transversality conditions that requires that the $\{v_t\}$ sequence remain bounded. The transversality condition will be used to pin down the correct solution of the Euler difference equations (8.11).

3. Factoring the Characteristic Matrix Polynomial Associated with the Euler Equation

We now indicate how the spectral-density-like polynomial $\{B'(LI - A')^{-1}R(L^{-1}I - A)B + Q\}$ can be factored and how this factorization permits obtaining the solution of the control problem in a convenient form.

Recall the equations for F and P:

$$F' = A'PB[B'PB + Q]^{-1} \tag{8.6}$$

$$P = A'PA + R - A'PB(B'PB + Q)^{-1}B'PA. \tag{8.7}$$

We now establish the following identity which gives an expression for the polynominal on the left side of (8.11):

Lemma 8.1: (Factorization Identity)'

$$\begin{aligned}&[I + B'(zI - A')^{-1}F'][Q + B'PB][I + F(z^{-1}I - A)^{-1}B] \\ &\qquad = Q + B'(zI - A')^{-1}R(z^{-1}I - A)B.\end{aligned} \tag{8.12}$$

Proof: First note the identity

$$P - A'PA = (zI - A')P(z^{-1}I - A) + A'P(z^{-1}I - A) + (zI - A')PA \tag{8.13}$$

To establish this identity, write out the right side to obtain

$$\begin{aligned}&P - zPA - A'Pz^{-1} + A'PA + A'Pz^{-1} - A'PA + zPA - A'PA \\ &\qquad = P - A'PA.\end{aligned}$$

Next, we substitute (8.7) for P into (8.13) to get

$$(zI - A')P(z^{-1} - A) + A'P(z^{-1}I - A) + (zI - A')PA + A'PB(B'PB + Q)^{-1}B'PA$$
$$= R.$$

Premultiply the above equation by $B'(zI - A')^{-1}$ and post multiply by $(z^{-1}I - A)^{-1}B$ to obtain

$$\begin{aligned} B'PB &+ B'(zI - A)^{-1}A'PB + B'PA(z^{-1}I - A)^{-1}B \\ &+ B'(zI - A)^{-1}A'PB(B'PB + Q)^{-1}B'PA(z^{-1}I - A)^{-1}B \\ &= B'(zI - A')^{-1}R(z^{-1}I - A)^{-1}B \end{aligned}$$

Now from (8.6) $A'PB = F'(B'PB + Q)$, which when substituted into the preceding equation gives

$$\begin{aligned} B'PB &+ B'(zI - A)^{-1}F'(B'PB + Q) + (B'PB + Q)F(z^{-1}I - A)^{-1}B \\ &+ B'(zI - A')^{-1}F'(B'PB + Q)F(z^{-1}I - A)^{-1}B \\ &= B'(zI - A')^{-1}R(z^{-1}I - A)B \end{aligned} \tag{8.14}$$

Notice that

$$\begin{aligned} [I + B'(zI - A)^{-1}F'](B'PB + Q)[I + F(z^{-1}I - A)^{-1}B] \\ = B'PB + Q + (B'PB + Q)F(z^{-1}I - A)^{-1}B \\ + B'(zI - A')^{-1}F'(B'PB + Q) \\ + B'(zI - A')^{-1}F'(B'PB + Q)z^{-1}I - A)^{-1}B \end{aligned}$$

In light of the above equality, adding Q to both sides of (8.14) gives the factorization identity

$$\begin{aligned} [I + B'(zI - A')^{-1}F'](B'PB + Q)[I + F(z^{-1}I - A)^{-1}B] \\ = Q + B'(zI - A')^{-1}R(z^{-1}I - A)B. \blacksquare \end{aligned} \tag{8.12}$$

The factorization identity (8.12) is a special case of another factorization identity that is associated with linear regulation problems in which there are cross products between states and controls in the objective function. We state this identity in the following lemma.

Lemma 8.2: (General Factorization Identity)

Let F and P satisfy

$$F = (Q + B'PB)^{-1}(B'PA + W') \tag{8.15}$$

$$
(8.16)\qquad \begin{aligned} P &= R + A'PA \\ &\quad - (A'PB + W)\,(Q + B'PB)^{-1}\,(B'PA + W'). \end{aligned}
$$

(These equations are the formula for the optimal feedback law and the algebraic matrix Ricatta equation, respectively, for a linear regulator with cross product between states and controls described by matrix W.) The following identity holds:

$$
(8.17)\qquad \begin{aligned} &Q + B'(z^{-1}I - A')^{-1}R(z^{-1}I - A)^{-1}B + B'(z^{-1}I - A')^{-1}W + W'(z^{-1}I - A)^{-1}B \\ &= [I + B'(zI - A')^{-1}F']\,(Q + B'PB)[I + F(z^{-1}I - A)^{-1}B]. \end{aligned}
$$

Proof: The proof precisely parallels the steps for proving lemma 8.1. ∎

We will not need to use the more general lemma 8.2 here, but it comes in handy sometimes (see hansen and Sargent [1988]).

We shall study the structure of the factorization (8.12) of the matrix polynomial associated with the Euler equation by using the following two facts from matrix algebra. Let a, b, c, and d be matrices and let all of the indicated inverses exist. Then we have the identities

$$
(8.18)\qquad [a - bd^{-1}c]^{-1} = a^{-1} - a^{-1}b[d - ca^{-1}b^{-1}ca^{1}
$$

and

$$
(8.19)\qquad det\, d \cdot det\,(a - bd^{-1}c) = det\, a \cdot det\,(d - ca^{-1}b).
$$

For proofs see Fortmann [] or Nobel and Daniel [p. 29, 210]. Using (8.18) with $a = I, -b = F', d = (zI - A'), c = F'$ we have

$$
(8.20)\qquad [I + B'(zI - A)^{-1}F']^{-1} = I - B'[zI - (A' - F'B')]^{-1}F'
$$

Next with $a = I, b = B', d = (zI - (A' - F'B')), c = F'$, apply (8.19) to get

$$
\begin{aligned} &det\,[I - B(zI - (A' - F'B'))^{-1}F'] \cdot det\,(I - (A' - F'B')) \\ &= det\,[zI - (A' - F'B') - F'B'] \end{aligned}
$$

or

$$
(8.21)\qquad det\,[I - B(zI - (A' - F'B'))^{-1}F' = \frac{det\,(zI - A')}{det\,[zI - (A' - F'B')]}
$$

Combining (8.21) with (8.20) gives

$$det\,[I + B'(zI - A)^{-1}F']^{-1} = \frac{det\,(I - A')}{det\,(zI - (A' - F'B'))}$$

or

$$det\,[I + B'(zI - A)^{-1}F'] = \frac{det\,(zI - A' - F'B')}{det\,(zI - A} \tag{8.22}$$

Equation (8.22) implies the following:

Lemma 8.3: The zeroes of $det\,[I + B'(zI - a)^{-1}F']$ equal the eigenvalues of $A' - F'B'$, which equal the eigenvalues of $(A - BF)$. If the pair (A, B) is stabilizable, the zeroes of $det\,[I + B'(zI - A)^{-1}F']$ all are less than unity in modulus.

4. Solving the Non-Stochastic Problem

Armed with these results, we now return to study the Euler equation (8.11). The factorization (8.13) permits the Euler equation to be written as

$$\{[I + B'(LI - A')^{-1}F'](B'PB + Q)[I + F(L^{-1}I - A)^{-1}B]\}v_t = R_{21}a_t \tag{8.23}$$

In effect, the transversality conditions require that to get the correct solution we must operate on both sides of (8.23) with the inverse of $[I + B'(LI - A')^{-1}F'](B'PB + Q)$, which gives

$$[I + F(L^{-1}I - A)^{-1}B]v_t = (Q + B'PB)^{-1}[I + B(LI - A')^{-1}F']R_{21}a_t$$

Substituting $x_t = (L^{-1}I - A)^{-1}Bv_t$ and equation (8.20) in the above gives

$$v_t = -Fx_t + (Q + B'PB)^{-1}[I - B'\Big[LI - (A' - F'B'))\Big]^{-1}F'R_{21}a_t$$

This can be written

$$\begin{aligned} v_t = &- Fx_t + (Q + B'PB)^{-1}R_{21}a_t - (Q + B'PB)^{-1}B'[LI \\ &- (A' - F'B')]^{-1}F'R_{21}a_t. \end{aligned} \tag{8.24}$$

In effect, (8.24) expresses v_t in terms of a "feedback" part, $-Fx_t$, and a "feedforward" part, as the remaining terms on the right side of (8.24) forms a weighted sum of current and future a_t's.

To proceed with the analysis, we use the following theorem from Kwakernaak and Sivan [] and Zadeh and Desoer [] :

Theorem 8.1: (Leverrier's algorithm)

Consider the constant $(n \times n)$ matrix G with characteristic polynomial

$$det\,(zI - G) = z^n + \alpha_{n-1} z^{n-1} + \ldots + \alpha_1 z + \alpha_0.$$

Then

$$(zI - G)^{-1} = \frac{1}{det\,(zI - G)} \sum_{i=1}^{n} z^{i-1} R_i$$

where the $n \times n$ matrices R_i are given by

$$R_i \sum_{j=i}^{n} \alpha_i G^{j-i}, \; i = i, 2, \ldots, n,$$

with $\alpha_n = 1$. The coefficients α_i and the matrices R_i can be obtained from

$$\begin{aligned} \alpha_n &= 1, R_n = I, \\ \alpha_{n-k} &= \frac{1}{k}\, tr\,(GR_{n-k+1}), k = 1, 2, \ldots, n. \\ R_{n-k} &= \alpha_{n-k} I + GR_{n-k+1}, k = 1,\ 2, \ldots, n\,. \\ R_0 &= 0. \end{aligned}$$

For a proof of the theorem, see Kwabernaak and Sivan [1972,].

Applying theorem 8.1 with $G = (A' - F'B')$ gives

$$[LI - (A' - F'B')]^{-1} = \frac{1}{det\,(LI - (A' - F'B'))} \sum_{i=1}^{n} R_i L^{i-1}$$

or

$$[LI - (A' - F'B')]^{-1} = \frac{1}{(L^n + \alpha_{n-1} L^{n-1} + \ldots + \alpha_1 L + \alpha_0)} \sum_{i=1}^{n} R_i L^{i-1}$$

Since the eigenvalues of $(A' - F'B')$ are less than unity in modulus, we have

$$[LI - (A' - F'B')]^{-1} = \frac{1}{(L - \mu_1)(L - \mu_2) \ldots (L - \mu_n)} \sum_{i=1}^{n} R_i L^{i-1}$$

where $| \mu_i | < 1$ and the μ_i's are zeroes of $\sum_{i=0}^{m} \alpha_i L^i$, which equal the eigenvalues of $(A - BF)$, which we assume are distinct. Multiplying the numerator and denominator of the right side of the preceding equation both by L^{-n} gives

$$(8.25) \quad [LI - (A' - F'B')]^{-1} = \frac{1}{(1 - \mu_1 L^{-1})(1 - \mu_2 L^{-1}) \ldots (1 - \mu_n L^{-1})} \sum_{i=1}^{n} L^{-(n+1-i)} R_i$$

Substituting (8.25) into (8.24) gives

$$v_t = -Fx_t + (Q + B'PB)^{-1} R_{21} a_t$$
$$(8.26) \qquad - (Q + B'PB)^{-1} B' \frac{\sum_{i=1}^{n} L^{-(n+1-i)} R_i F' R_{21} a_t}{(1 - \mu_1 L^{-1})(1 - \mu_2 L^{-1}) \dots (1 - \mu_n L^{-1})}$$

The last term can be expressed as follows using matrix partial fractions:

$$(8.27) \qquad \frac{(Q + B'PB)^{-1} B' \sum_{i=1}^{n} L^{-(n+1-i)} R_i F' R_{21}}{(1 - \mu_1 L^{-1}) \dots (1 - \mu_n L^{-1})} = \sum_{j=1}^{n} \frac{C_j L^{-1}}{(1 - \mu_j L^{-1})}$$

where

$$C_j = \frac{(Q + B'PB)^{-1} B' \left[\sum_{i=1}^{n} (\mu_j)^{-(n-i)} R_i\right] F' R_{21}}{\prod_{k \neq j} (1 - \mu_k \mu_j^{-1})}$$

With (8.27) substituted into (8.26) the decision rule becomes

$$v_t = -Fx_t + (Q + B'PB)^{-1} R_{21} a_t$$
$$(8.28) \qquad - \sum_{j=1}^{n} \frac{C_j}{(1 - \mu_j L^{-1})} a_{t+1}$$

or

$$v_t = -Fx_t + (Q + B'PB)^{-1} R_{21} a_t$$
$$(8.29) \qquad \sum_{j=1}^{n} C_j \sum_{k=0}^{\infty} \mu_j^k a_{t+1+k}$$

This is the optimal plan for setting v_t when the a_t sequence is known with certainty. The reason for calling $(-Fx_t)$ the "feedback" part of the solution and the remaining part of the right side of (8.29) the "feedforward" part is now clear.

5. Solution Under Uncertainty

Under uncertainty about future a_t's, the appropriate solution is:

$$v_t = -\, Fx_t + (Q + B'PB)^{-1} R_{21} a_t$$
$$(8.30) \qquad \sum_{j=1}^{n} C_j \sum_{k=0}^{\infty} \mu_j^k E_t a_{t+1+k}$$

where E_t is the conditional expectation unconditioned on information known at t, i.e., $E_t(\cdot) = E(\cdot \mid z_t, z_{t-1}, \dots)$. An explicit formula for v_t in terms of the current and past z_t's can be derived using the procedures of Hansen and Sargent [, appendix A]. In particular, recall that we have assumed that a_t is a component of z_t, say $a_t = ez_t$ where e is a vector linking a and z, e.g. often a vector of zeroes and ones. We have assumed that

$$\rho(L)z_t = v_t^z$$

or

$$z_t = \xi(L)v_t^z \equiv \rho(L)^{-1}v_t^z$$

where our assumption that the zeroes of $det\,\rho(z)$ are outside the unit circle quarantee that $\rho(L)^{-1}$ is one-sided and square summable in nonnegative powers of L. We want to form terms of the form

$$\sum_{k=0}^{\infty} \mu_j^k\, E_t\, a_{t+1+k}$$

Using the Wiener-Kolomogorov theory of prediction and the results of Hansen and Sargent [1980], we have

$$\sum_{k=0}^{\infty} \mu_j E_t\, a_{t+1+k} = e[\frac{L^{-1}\xi(L)}{1-\mu_j\, L_{-1}}]_+ v_t^z \tag{8.31}$$

where the operator $[\ \ ]_+$ means "ignore negative powers of L", i.e. $[\sum_{j=-\infty}^{\infty} h_j L^j]_+ = \sum_{j=0}^{\infty} h_j L^j$. Using the technique of Hansen and Sargent, the right side of (8.31) can be shown to be

$$\begin{aligned} e[\frac{L^{-1}\xi(L)}{1-M_j L_{-1}}]_+ v_t^z &= e[\frac{L^{-1}-L^{-1}\xi(\mu_j)\xi^{-1}(L)}{1-\mu_j L^{-1}}]_+ z_t \\ &= e\rho(\mu_j)^{-1}\Big[\sum_{j=1}^{r-1}\Big(\sum_{k=j+1}^{r}(\mu_j)^{k-j-1}\rho_k\Big)L^j\Big]z_t, \end{aligned}$$

where recall that r is the order of the autoregressive process for z_t. Substituting the above into the right side of (8.30) gives

$$\begin{aligned} v_t = &-Fx_t + (Q+BPB)^{-1}R_{21}a_t \\ &- \Big\{\sum_{j=1}^{n} C_j\, e\rho(\mu_j)^{-1}\Big[\sum_{j=1}^{r-1}(\sum_{k=j+1}^{r}(\mu_j)^{k-j-1}\rho_k)L^j\Big]\Big\}z_t \end{aligned} \tag{8.32}$$

$$\rho(L)_t = v_t \tag{8.31}$$

Equations (8.32) and (8.31) compactly display the restrictions across the $\{z_t\}$ process and the optimal control law for $\{v_t\}$ that are implied by the dynamic optimum theory.

Sometimes optimum theory problems are encountered with criterion functions of the form (P1), except that the term $-2a_t' R_{21} v_t$ is replaced by a term of the form $-2a_t' R_{21} x_t$. With minor modifications, the preceding solution applies in this case. Consider the term

$$J = \sum_{t=t_0}^{\infty} \bar{a}_t' \bar{R}_{21} x_t = \sum_{t=t_0}^{\infty} \bar{a}_t' \bar{R}_{21} (L^{-1} I - A)^{-1} B v_t.$$

Writing out this last sum and differentiating J with respect to v_t, one finds

$$\frac{\partial J}{\partial v_t} = B'(LI - A)^{-1} \bar{R}_{21} \bar{a}_t$$

The Euler equation thus becomes (8.11) with $B'(LI - A)^{-1} \bar{R}_{21} \bar{a}_t$ replacing $R_{21} a_t$ on the right side. An equivalent procedure is to define

$$(R_{21} a_t)' = (B'(LI - A)^{-1} \bar{R}_{21} \bar{a}_t)' \tag{8.33}$$

With this definition of $R_{21} a_t$, the Euler equations for the amended problem remain (8.11).